



Aplicaciones e innovación de la ingeniería en ciencia y tecnología

Esteban Inga Ortega (Coord.)

INGA ORTEGA, E., ed. *Aplicaciones e innovación de la ingeniería en ciencia y tecnología* [online]. Quito: Editorial Abya-Yala, 2019, 264 p. ISBN: 978-9978-10-491-0. https://doi.org/10.7476/9789978104910.



Aplicaciones e innovación de la ingeniería en ciencia y tecnología

Esteban Inga Ortega
Coordinador

Aplicaciones e innovación de la ingeniería en ciencia y tecnología



2019

Aplicaciones e innovación de la ingeniería en ciencia y tecnología

1era. Edición

Ediciones Abya-Yala
Av. 12 de octubre N 23-116 y Wilson
Casilla 17-12-719
Telf.: (593-2) 2506251
Fax: (593-2) 2506267
E-mail: editorial@abyayala.org
www.abyayala.org
Quito-Ecuador

CARRERA DE ELECTRICIDAD
CARRERA DE ELECTRÓNICA Y AUTOMATIZACIÓN
CARRERA DE INGENIERÍA CIVIL
CARRERA DE COMPUTACIÓN
CARRERA DE TELECOMUNICACIONES
CARRERA DE BIOMEDICINA

Derechos de autor: 055809
ISBN: 978-9978-10-361-6

Tiraje: 300 ejemplares

Ilustración de portada: https://www.shutterstock.com

Diseño, Diagramación e Impresión: Editorial Abya-Yala
Quito-Ecuador

Impreso en Quito-Ecuador, junio 2019

Publicación arbitrada de la Universidad Politécnica Salesiana

Índice

Capítulo IV

Capítulo V

Capítulo VI

Capítulo VII

Prólogo

Es un placer introducir esta colección de artículos científicos del área de Ciencia y Tecnología de la Universidad Politécnica Salesiana, trabajo presentado desde las ingenierías: Electricidad, Electrónica y Automatización, Civil, Computación, Telecomunicaciones y Biomedicina.

El mundo ha avanzado con la llegada de la ciencia y tecnología. Satisfacer las necesidades de la sociedad, con una visión de innovación, se ha convertido en un reto para el campo científico.

El camino para llegar a un concepto de ciudades inteligentes, por ejemplo, puede conjugar varias aristas que dan cuenta de un aporte de diversas competencias y destrezas por parte de la comunidad científica; de esta manera, podemos encontrar aportes en redes eléctricas inteligentes, servicios de comunicación masiva, aprovechamiento de los recursos hídricos, análisis de ondas sísmicas, manejo de datos en la nube o la interpretación de imagen para aplicaciones médicas, cumpliendo así una vasta demanda de oportunidades para la generación de nuevo conocimiento que aporta la ciencia y tecnología en favor de la sociedad.

Este libro presenta los resultados y esfuerzos en investigación básica y aplicada en un contexto articulado en Aplicaciones e innovación de ingeniería en ciencia y tecnología, relacionados con aspectos importantes en el área de energía, electrónica, civil, computación, telecomunicaciones y biomedicina.

El trabajo incluye documentos seleccionados sobre la base de la originalidad, importancia y claridad en el tema. Esperamos que personas de la academia y de la industria encuentren en este libro herramientas para el desarrollo de nuevas investigaciones sobre los aportes que se pueden conseguir desde la ingeniería contribuyendo con la sociedad, la ciencia y la tecnología.

Este trabajo no se podría haber llevado a cabo sin el apoyo de muchas personas de la Universidad Politécnica Salesiana que siem-

pre están en búsqueda del desarrollo del conocimiento como un bien de uso común.

Introducción

Las carreras de ingenierías de la Universidad Politécnica Salesiana Sede Quito en conformidad con las necesidades expuestas en el plan nacional de desarrollo en relación con el contexto y necesidades propias de la sociedad ecuatoriana y mundial, realizan una exposición de los avances científicos en ciencia y tecnología brindando un aporte basado en investigación básica y aplicada con el que se da solución a problemas de ingeniería con aporte y responsabilidad social.

A continuación, se esbozan los contenidos que abarca este libro:

Capítulo I: Desarrolla un modelo matemático heurístico para despliegue de redes inalámbricas heterogéneas que brindarán la conectividad de los medidores inteligentes de energía eléctrica; adicionalmente el trabajo incluye el despliegue de una red de distribución eléctrica; en ambos casos se minimizan los costos por el uso de recursos tanto para la red de comunicaciones inalámbrica como para la red eléctrica.

Capítulo II: Presenta un modelo para localización de fallas en sistemas de distribución eléctrica que permita reducir las pérdidas por energía no facturada; además con este modelo se busca aumentar la confiabilidad del sistema. El trabajo entonces presenta una revisión de métodos para localizar fallas en los sistemas eléctricos de distribución eléctrica con presencia de fuentes de generación distribuida.

Capítulo III: Presenta una metodología para cuantificar la intercepción de neblina a través de un sistema de captación y evaluar su potencial como fuente de agua. Se trabaja con tres tipos de colectores de neblina. Se analizan las variables meteorológicas con el que cada colector se relaciona para un rendimiento mayor del colector de acuerdo a las condiciones del clima.

Capítulo IV: Presenta una metodología para estudios de sitio para determinar la degradación de rigidez y resistencia de los suelos, la evaluación del potencial de licuación de arenas, niveles de

deformación en los estratos de suelos y el espectro de diseño para su posterior uso en el análisis dinámico de estructuras. De esta manera, se pueden construir estructuras dúctiles que pueden experimentar grandes deformaciones sin pérdida de resistencia para que no colapsen ante un sismo severo.

Capítulo V: Se presenta una arquitectura destinada a integrar sistemas embebidos FPGA que poseen menor consumo de energía como acelerador en un Cloud privado. Para verificar su rendimiento se tiene en cuenta la calidad del servicio logrando una combinación adecuada entre CPU y FPGA en un entorno Cloud.

Capítulo VI: Se advierte un despliegue de una red de fibra óptica en una zona urbana para un número considerable de usuarios del tipo residencial y corporativo y así satisfacer la demanda por ancho de banda. El modelo considera escalabilidad y una inversión redituable. El modelo considera distintos tipos y estándares de redes de fibra óptica.

Capítulo VII: Se presenta un aporte desde la Telemedicina en la cual se considera el tratamiento de imágenes para un correcto diagnóstico médico desde un análisis visual. Se presenta un esquema de clasificación automática de nódulos para diferentes proyecciones mamográficas. El modelo propuesto se encarga de fusionar la información y clasificar nódulos malignos y benignos. Información adecuada para los profesionales encargados de radiología.

Capítulo I

Óptima planeación de recursos para redes eléctricas inteligentes usando modelos heurísticos

Optimal Planning of Resources for Smart Grid using Heuristic Methods

Esteban Inga[*], Miguel Campaña[§]
Universidad Politécnica Salesiana Ecuador
[*]Email: einga@ups.edu.ec
[§]Email: mcampana@ups.edu.ec

Resumen

El presente trabajo considera una planificación de redes inalámbricas heterogéneas para proporcionar medición inteligente dentro de una red de área vecina con su respectiva conectividad con la oficina central de una compañía de distribución eléctrica. Además, el trabajo presenta un modelo de planificación para lograr una red de distribución eléctrica subterránea en un área definida. Tanto la red inalámbrica heterogénea como la red de distribución se implementan en un área georeferenciada real y permiten minimizar el costo por el uso de recursos de telecomunicaciones y eléctricos; de esta forma, se logra el concepto de una red eléctrica inteligente eficiente al minimizar los costos y aumentar la confiabilidad de la red considerando las restricciones de capacidad y la cobertura del usuario.

Palabras claves

Sistemas eléctrico de distribución; optimización; planeación; medición inteligente; redes eléctricas inteligentes

Abstract

The present work considers a planning of heterogeneous wireless networks to provide smart metering within a neighborhood area network with its respective connectivity with the central office of an electricity distribution company. In addition, a planning model is present to achieve an underground electrical distribution network in a defined area. Both

the wireless heterogeneous network and the distribution network are deployment over a real georeferenced area and allow minimizing the cost by use of telecommunications and electrical resources; in this way, the concept of an efficient intelligent electric network is meet by minimizing costs and increasing reliability of the network considering capacity restrictions and user coverage.

Keywords

Electrical distribution system; optimization; planning; smart metering; smart grid.

1. Introducción

Para facilitar la conversión de una red eléctrica convencional a una red eléctrica inteligente es necesario incluir un sistema de comunicaciones que permitan el intercambio de información bidireccional [1],[2]. El presente trabajo se enfoca en la infraestructura de medición inteligente de energía eléctrica; para lo cual, se requiere la conectividad entre el medidor inteligente de cada usuario y la oficina central de una empresa eléctrica, logrando una medición inteligente del suministro eléctrico que se provee a cada usuario o cliente por parte de una empresa de distribución eléctrica [3].

La infraestructura de medición inteligente podrá estar formada de varias subredes con diferentes tecnologías de comunicación para cada caso, así tenemos que la red de área doméstica estará dedicada al intercambio de información entre el medidor inteligente y los sensores ubicados en el interior de la vivienda, esta información podrá ser utilizada para caracterizar el consumo de cada usuario de manera particular, en este punto, un detalle importante se relaciona a la seguridad de la información, que aunque no está involucrado en el presente trabajo, debe preocupar a las empresas eléctricas y a los usuarios, ya que, podría sufrir ataques de parte de la delincuencia y con este conocer el estado y hasta el número de dispositivos electrónicos que posee un cliente, tiempo de uso, tarifa de pago, etc., [4], [5].

A continuación, se desprende la red de área vecindaria, pilar fundamental de este trabajo, donde se encuentra la mayor necesidad de optimizar los recursos, ya que, los usuarios podrán estar ubicados en zonas urbanas con mayor cercanía entre una red de comunicaciones; existente en relación con una zona suburbana, pero también estarán las zonas rurales con baja penetración de la

red de comunicaciones y requerirá que los recursos de tecnología de comunicaciones sea la idónea. En todos los casos anteriores se deberá tener una planeación adecuada que permita una rápida instalación de medidores inteligentes, pero de forma escalable; de esta manera se garantiza que el costo por recursos sea el mínimo. Los criterios de partida para iniciar el despliegue no solo podrán depender de atacar zonas donde existe mayor robo o hurto de energía eléctrica, sino que, a qué tipo de zona se refiere y qué característica en disponibilidad de tecnología de comunicaciones existe para determinar el hardware que deberá disponer un medidor inteligente. De lo antes mencionado, es prescindible resaltar que una única solución de comunicaciones no es viable, sino que, puede existir múltiples soluciones y con el menor costo [6],[7].

La red de área amplia encargada de la conectividad entre los medidores inteligentes con la empresa de distribución eléctrica pasa a formar una zona de frontera debido al requerimiento de concentradores que aglutinarán información proveniente de un conjunto de medidores inteligentes y a su vez permite el intercambio de información o enlace de última milla con una estación base celular, elemento que finalmente enlaza o envía información con la oficina central [8].

Tal como se aprecia el envío de información desde un medidor inteligente a una empresa de distribución eléctrica toma el nombre de enlace ascendente, pero se debe tener presente que la información puede ser descendente debido a que la medición inteligente también permite cortes y reconexiones del suministro eléctrico con lo que a través de la información recibida se activa o desactiva el sistema electrónico del medidor inteligente y así se permite o no el paso del suministro eléctrico a la vivienda del cliente [9],[10].Como complemento a la planeación de la red de comunicaciones para medición inteligente es primordial dotar de un modelo que permita, en la misma zona georreferenciada, el dimensionamiento de una red eléctrica capaz de proporcionar mapas de rutas óptimos con la finalidad de construir redes eléctricas, ya sean: soterradas, aéreas o en su defecto mixtas muy pertinentes para facilitar que la red eléctrica inteligente consiga confiabilidad y calidad del suministro [11]–[15].

La creciente demanda por el incremento de cargas significativas a la red eléctrica de distribución existente, como: introducción

de cocinas de inducción, calentamiento de agua, incremento de nuevos usuarios a la red eléctrica o por la integración de vehículos eléctricos en un inmediato próximo [16] hace que, la necesidad de redimensionar equipos eléctricos de potencia reconsiderando capacidades nominales de corriente en los conductores y otras variables eléctricas sensibles al incremento de la demanda sean nuevamente analizados para garantizar un servicio eléctrico de calidad y confiable [17],[18]. En redes eléctricas nuevas, se debe planificar considerando las demandas tipo actualizadas que contemplan la energía demandada por cocción y calentamiento de agua. Por lo tanto, se propone implementar una heurística capaz de solucionar el problema de planeación en redes eléctricas de distribución considerando restricciones de cobertura lineal y capacidad de albergar usuarios en los transformadores de distribución a emplazarse en sitios georreferenciados, determinados por la heurística, garantizando satisfacer el máximo porcentaje de observabilidad a los usuarios sin descuidar los principios técnicos de calidad y seguridad energética. El modelo será implementado en el software Matlab y mediante el software Cymdist se analizará el flujo de potencia resultante obtenido del mapa de ruta lograda por la heurística propuesta.

El modelo de planeación de redes de distribución es capaz de proporcionar datos de mucho interés para iniciar un despliegue óptimo en una red eléctrica de distribución nueva, tales como: cantidad de transformadores por cada capacidad en kVAs requerido en el despliegue, caídas de voltaje máximos en cada transformador emplazado y capacidad máxima requerida en la subestación para satisfacer la demanda total. Además, el modelo admite parámetros reales de diseño, tales como: kVA-m soportado por el conductor a utilizarse en bajo voltaje, demandas por cada usuario identificando a qué estrato pertenece, estandarización de transformadores y distancia promedio en la que se debe emplazar los pozos en una red eléctrica soterrada o los postes en una red eléctrica aérea. El modelo da solución al problema de planeación aplicando teoría de grafos y enrutamiento de mínima y máxima expansión [19], garantizando así, proporcionar topologías al menor costo.

El análisis eléctrico obtenido del software Cymdist, tomado en cuenta, el despliegue y datos obtenidos por la heurística propuesta conocido en la literatura como modelo de red de referencial [20],

contribuirá de manera óptima a la toma de decisiones garantizando una red eléctrica segura y resiliente al menor costo de implementación. El algoritmo de planeación y despliegue de una red eléctrica de distribución consta de cinco procedimientos. El primer procedimiento consiste en levantar la información necesaria georreferenciada de un archivo OSM obtenido de OpenStreetMap, el mismo que, contiene la información necesaria de las coordenadas (latitud y longitud) de calles, viviendas, parques, etc.

En el segundo procedimiento se genera múltiples escenarios de conectividad observando las restricciones anteriormente mencionadas con técnicas de clusterización basado en una variante del algoritmo de PRIM. Mediante el algoritmo del Setcover se procede, en el paso tres, a seleccionar el mínimo número de subconjuntos de conectividad que contengan el mayor número de elementos, para así, proporcionar una solución cercana a la óptima.

En consecuencia, la solución obtenida por la heurística, no se basa en la topología de una red inicial, sino que, se basa en la conjugación de varias soluciones, de tal manera que, se busca garantizar una solución óptima global.

En el procedimiento cuatro, se eliminan elementos redundantes de los subconjuntos seleccionados por el algoritmo Setcover, posteriormente, de cada subconjunto o agrupación, bajo el criterio del centro de masa, se selecciona la ubicación del transformador a emplazarse y se procede a construir un árbol de mínima expansión, mediante el algoritmo de Dijkstra entre los primarios de los transformadores habilitados dando origen a la red de medio voltaje y el mismo criterio de aplica en los secundarios, con cada agrupación, para formar la red de bajo voltaje, a diferencia que, se emplea un árbol de máxima expansión con la finalidad de considerar a todas las viviendas de cada subconjunto para el tendido eléctrico desde los secundarios de los transformadores de distribución hacia sus nodos de servicio más lejanos. Por último, se procede a calcular las caídas de voltaje en cada transformador emplazado (desde la fuente hacia el nodo más lejano).

En adelante este artículo se organiza de la siguiente manera. En la sección 2 se introduce el modelo de planeación para redes de comunicaciones inteligente y para redes eléctricas de distribución. En la sección 3 se describe la formulación del problema para una red

heterogénea inalámbrica y la red eléctrica de distribución soterrada. En la sección 4 se analizan los resultados de la modelización y simulación de los algoritmos de planeación. Finalmente, concluimos este artículo en la sección 5.

2. Planeación Óptima de Redes

A. Red heterogénea inalámbrica para medición inteligente

Si consideramos medidores inteligentes de energía eléctrica con intervalos de lectura cada 15 minutos, en 24 horas tendríamos 96 lecturas por día y en un mes 2880 lecturas. Por otro lado, si se considera que cada lectura representa 5 bytes; en un escenario de 4'000.000 de abonados, al multiplicar por 5 bytes nos genera una cantidad aproximada a 20'000.000 bytes por lectura. La Tabla 1, muestra el crecimiento de la información para un intervalo entre cada lectura [21],[11], [22]. El concepto de los recursos requeridos por la infraestructura de medición avanzada se avizora en la Figura 1, de acuerdo a este modelo se incorporan concentradores o puntos de agregación de datos (PAD) que permitirán la conectividad con las estaciones base celular y de esta manera se reducen los costos al incorporar tecnología inalámbrica con un costo menor a la tecnología celular.

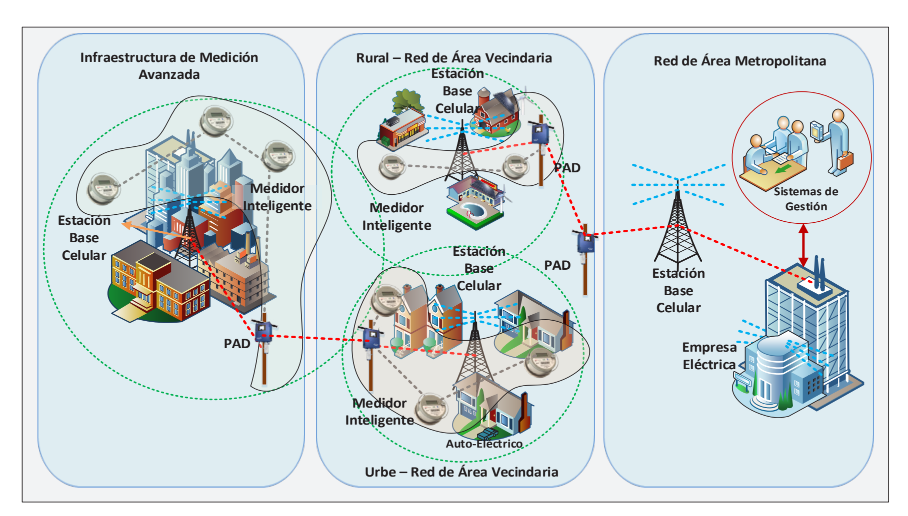


Figura 1. Infraestructura de medición avanzada usando redes inalámbricas

Tabla 1. Cantidad de información para un intervalo de lectura de 15 minutos

Número de Medidores	Datos	Lecturas	Intervalo de Tiempo	Cantidad de Datos
4M	5 Bytes	96	1 Día	1.92GB
4M	5 Bytes	96	1 Mes	57.6GB
4M	5 Bytes	96	1 Año	691.2GB

De los datos preliminares se puede apreciar que el crecimiento de la información es considerable, más aún si se requieren obtener reportes de esa cantidad de datos de manera eficiente para una posterior toma de decisiones. Es así que, lograr una infraestructura de medición avanzada, resulta importante investigar para lograr minimizar el impacto que esta red dispondría para una empresa de distribución eléctrica.

Para crear una infraestructura de medición avanzada se han realizado varios aportes científicos que han buscado dar solución desde diferentes frentes y así conseguir una medición inteligente de energía eléctrica; sin embargo, las restricciones propias de la tecnología, las zonas, el costo, hacen de la planeación de la red de comunicaciones un problema no trivial que en términos de optimización se describe como un problema NP-Complete dentro de los capítulos propios de problemas combinatorios; por lo tanto, conseguir una única solución que nos entregue el óptimo global de la planeación de la red de comunicaciones resulta inviable; sin embargo, es posible advertir que las tecnologías se encuentran maduras y/o ya desplegadas con el objeto de minimizar el costo y garantizar la confiabilidad de la red [2], [5], [23]–[25].

En sus inicios para conseguir una medición inteligente se avizoraba un crecimiento oportuno para una solución con radio frecuencia ya sea en el rango de 902-928 MHz o en 169MHz, en este sentido, aunque no existía costo por uso de frecuencia sí se requería un número considerable de concentradores con RFID; finalmente esta solución no dio eficiencia en el despliegue debido a que la configuración y mantenimiento de los recursos los debía absorber las empresas de distribución eléctrica destinadas tradicionalmente a un trabajo diferente por la naturaleza de su negocio. En un camino por solventar el problema de cobertura y permitir la bidireccionalidad de la red, se

realizaron enlaces de última milla colocando una tarjeta de comunicación con tecnología celular que permita el enlace con las estaciones base celular más cercanas.

Posterior a los inicios de RFID orientado a disminuir las pérdidas no técnicas de energía proveniente del robo o hurto de clientes residenciales, se advirtió una posibilidad para la medición inteligente de energía eléctrica de clientes industriales o especiales ubicados de manera aislada pero dentro de zonas con cobertura celular, dando origen a una solución única con tecnología celular, este costo por los recursos y el servicio del operador primario de telefonía celular al ser un escenario pequeño no presentaba dificultad y se generaron dichos despliegues con satisfacción compartida, tanto para el cliente industrial,como para la empresa de distribución que entrega el servicio eléctrico.

Actualmente, operadoras de telefonía celular han intentado dar servicio de medición inteligente incluyendo opciones para energía eléctrica, agua y gas, basándose en que únicamente cambia el medidor inteligente y han realizado pruebas NB-IoT que utilizan bandas celulares y que nació tradicionalmente como una tecnología de radio de banda estrecha para aplicaciones de internet de las cosas, por otro lado, se tiene el uso del ahora estándar IEEE 802.15.4g que tiene un costo menor a WiFi y aunque ha mejorado el alcance no ha llegado a destacar en aplicaciones de medición inteligente.

Finalmente, y con mayor penetración en soluciones de medición inteligente de energía eléctrica se encuentra WiFi utilizada en exteriores que puede lograr conectividad entre concentradores y medidores inteligentes; por lo tanto, puede convertirse en una fuerte opción para lograr una planeación escalable de medición inteligentes de energía eléctrica, agua y gas. Para lograr que esta solución sea factible de implementar se debe tener en cuenta que al ser una solución híbrida se permite acoplar varias tecnologías, en este caso, de índole inalámbrico para formar una red heterogénea la misma que brinda la posibilidad de realizar multi saltos entre medidores inteligentes en zonas urbanas donde existen viviendas circundantes en un radio pequeño y en zonas suburbanas y rurales donde no exista posibilidad de multi salto entre el hardware de comunicación de cada medidor inteligente [26],[27]–[31].

B. Red de distribución eléctrica soterrada

La necesidad de actualización de las redes eléctricas de distribución debido a la creciente penetración de recursos energéticos distribuidos, junto con el crecimiento de la carga, eficiencia energética, seguridad y calidad del suministro eléctrico han cambiado de forma significativa los enfoques tradicionales de planeación, diseño y operación del sistema de energía eléctrica [32]. Para superar los desafíos operacionales del sistema eléctrico de distribución se han empleado un gran número de algoritmos, siendo implementados en unas pocas redes específicas de prueba que aún no son capaces de proporcionar conclusiones sólidas y generalizadas [33],[34],[35].

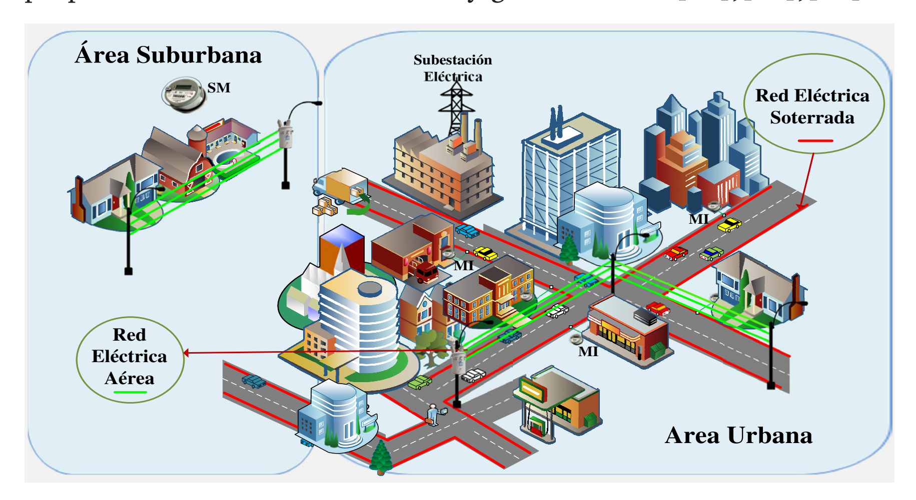


Figura 2. Planeación de redes eléctricas de distribución considerando arquitecturas soterradas y aéreas

En el presente documento se analiza desde dos enfoques realistas que son propiedades intrínsecas en los sistemas eléctricos de distribución, los cuales son: propiedades topológicas y propiedades eléctricas. Ambas propiedades eléctricas y topológicas de redes eléctricas de distribución tienen un impacto significativo en el rendimiento global de la red incluyendo caídas de tensión, pérdidas de potencia, fiabilidad de la red y, por último, cada impacto se ve reflejado en los costos. Por lo tanto, la presente investigación se centra en la planificación de redes eléctricas de distribución considerando restricciones reales con

la finalidad de proporcionar topologías seguras para la implementación al menor costo como se ilustra en la Figura 2 y, por último, en un segundo escenario, mediante el software Cymdist, trabajar sobre la segunda propiedad de una red eléctrica de distribución.

Con la finalidad de comprender y modelar una red compleja real soterrada y con la posibilidad de ser aérea o, en su defecto mixta, un conjunto de herramientas y técnicas son necesarias. En consecuencia, estas técnicas y herramientas tienen que ver con análisis de redes complejas y teoría de grafos. La red eléctrica soterrada dispone de varios elementos tales como líneas de transmisión, transformadores de distribución, pozos, zanjas y dispositivos de protecciones y maniobra eléctrica. Un modelo gráfico puede ser fácilmente construido teniendo en cuenta la relación de cada uno de los elementos existentes en la red eléctrica a desplegarse.

Para poder construir redes eléctricas de distribución es de mucha importancia considerar definiciones que se utilizan en la investigación topológica, tales como: propiedades de un grafo, propiedad de profundidad y propiedad topológica en redes eléctricas de distribución considerando coordenadas georreferenciadas.

Un grafo está formado por vértices (V) y aristas (E), donde, su expresión formal es $G=(V,E)$. En redes eléctricas de distribución, los vértices están asociados a los trasformadores de distribución, subestaciones, usuarios, interruptores, barras colectoras y pozos (redes soterradas) o postes (redes aéreas). La información de cada elemento, en relación con su posición, está dada en coordenadas geolocalizadas (Latitud y Longitud), donde, i representa el vértice de origen y j el vértice destino, que, mediante una arista E ambos vértices se relacionan. Al gráfico de conectividad se lo representa con una matriz de adyacencia cuadrada nxn, la misma que, relaciona a todos los vértices con sus respectivas aristas de manera controlada, es decir, sujeta a restricciones. Dichas restricciones en el presente documento son distancias lineales y capacidad de los trasformadores de distribución de albergar usuarios. Por lo tanto, la matriz cuadrada de conectividad es una matriz binaria, donde, 1 representa la existencia de una arista de un nodo vi a un vj y, por el contrario, 0 representa la no conectividad o la no existencia de una arista de un vi a un vj. En consecuencia, la matriz de conectividad se irá llenando de valores binarios coherentemente con las restriccio-

nes consideradas en el modelo para dar solución a la planeación de redes eléctricas de distribución. Por último, dentro de las propiedades del grafo, se considera, la longitud de la red que está dada por las distancias acumuladas o parciales de las aristas, además se considera la agrupación de los nodos que están directamente asociados con la capacidad del transformador de distribución.

En la propiedad de profundidad se analiza la topología armada por el grafo no dirigido *G* desde la fuente hacia los usuarios finales, el mismo que, en redes eléctricas de distribución en los niveles de medio y bajo voltaje son radiales o tipo árbol [32]. Un grafo con topología tipo árbol está definido por la utilización mínima de aristas para la conexión de todos sus vértices, es decir, si es el número de nodos o usuarios un árbol relaciona a todos sus vértices con *n-1* aristas. En el presente documento se genera la ruta al mínimo costo utilizando el algoritmo de Dijkstra en la red de medio voltaje y mediante una breve modificación se genera el árbol de máximo costo para la red de bajo voltaje. En la red de bajo voltaje se maximiza, puesto que, es necesario satisfacer de cobertura a todos los usuarios relacionados con sus respectivos subconjuntos.

En las propiedades topológicas de la red de distribución se observa la realidad del escenario, en el que, se contempla la cantidad de nodos por km^2, longitud de la red en medio voltaje requerido por km^2, número de alimentadores y número de pozos en el caso de despliegues soterrados, son variables que, claramente caracterizan las diferencias topológicas entre tipos de redes aéreas o soterradas. Otro elemento determinante para la topología de la red eléctrica es la técnica aplicada de clusterización, la misma que, agrupa elementos de objetos similares. En la presente investigación empleamos para dicho cometido, una variante del algoritmo de PRIM, sin considerar, el conocimiento de la cantidad de agrupaciones existentes, es decir, el modelo determina el número óptimo de agrupaciones para garantizar el máximo número de usuarios cubiertos por la red eléctrica.

3. Formulación del problema

Esta sección está dividida en dos partes: a) describe las variables que describen el problema para la planeación de una red heterogénea inalámbrica dedicada a la medición inteligente de energía eléctrica y b)

las variables que se involucran para resolver el problema de ubicación de transformadores en una red de distribución eléctrica soterrada.

A. *Red heterogénea inalámbrica*

Si partimos de formalismos matemáticos, el problema de ubicación óptima de PADs en una red de área vecindaria se la puede describir desde el concepto de teoría de grafos como un grafo no dirigido, el mismo que describe en un conjunto de vértices y como los enlaces, en este trabajo los medidores inteligentes se encuentran representados por los vértices y la red inalámbricas por los enlaces.

El escenario propuesto plantea tres elementos o recursos, estaciones base celular, PADs y MIs; para ello, es necesario tener en cuenta que un MI puede enviar información a un PAD o a otro MI lo que permite que a través de multi salto se envíe información hasta un PAD. Existe conectividad entre PADs para lograr conectividad con la estación base celular más cercana que finalmente realizar el trabajo de última milla con la empresa de distribución eléctrica.

Para el escenario propuesto los enlaces dentro del grafo no dirigido representan el costo determinado para este caso por la distancia, entonces se consideran únicamente valores positivos. Se supone que la tecnología inalámbrica a ser utilizada tanto en los MIs como en PAD es flexible, pero para el escenario se considera una red WiFi y celular. De esta manera el radio de transmisión puede ajustarse a las necesidades de la zona de despliegue, determinadas por las zonas urbanas, suburbanas o rurales; entonces, no todos los MIs tienen una única tecnología de comunicación. El alcance de comunicación de cada MI lo describe la variable ; además, definamos un conjunto $U= \{s1, s2, \dots sN\}$, de MIs. La posición del MI i-ésimo está dada por (xs_i, ys_i), donde $xsi,ysi \in S$.. El MI de coordenadas xsi,ysi puede conectarse a otro MI de coordenadas $xsj,ysj\ s$ si la distancia es menor o igual a distancia permitida $|(xs_i,ys_i)-(xs_j,ys_j)| \leq dmin$. Por lo tanto, se tiene un subconjunto $S_{dmin}\ (xui,yui)$, así todo MI que tenga alcance en el rango de la tecnología inalámbrica propuesta está dada por $S_{dmin}\ (xu_i,yu_i) = |(xs_i,ys_i)-(xs_j,ys_j)| \leq dmin,\ \forall_j \in S$

En caso de que un MI no esté en el rango de conectividad de otro MI podrá conectarse a un PAD, para este caso se ha restringido la posibilidad que un MI pueda conectarse directamente a una esta-

ción base celular con el objeto de reducir los costos; sin embargo, un MI puede quedar sin conectividad ya sea por capacidad de MI más cercano o del PAD, pero de estar en el radio de alcance de una estación base celular puede ser conectado por tecnología celular directamente. Es importante recordar que un MI en su arquitectura tiene la flexibilidad de contar con un módulo de tecnología inalámbrica WiFi, IEEE 802.15.4g, Celular, etc. Así, tenemos un grafo que garantiza la conectividad de los tres recursos inicialmente expuestos.

Para que un modelo de planeación sea considerado para uso real en las acciones de despliegue por parte de una empresa de distribución eléctrica es necesario contar con coordenadas georreferenciadas de latitud y longitud; además, de un cálculo de la distancia con la fórmula de haversine que indica la ecuación 1, fórmula que considera la curvatura de la tierra. De esta manera xs_j, ys_i es la longitud y latitud de los MIs

$$Distancia = 2 * R * asin^2 \sqrt{sin^2\left(\frac{\Delta lat}{2}\right) + cos(lat1)*cos(lat2)* \sin^2\left(\frac{\Delta lon}{2}\right)} \quad (1)$$

Para aplicar la ecuación 1 se debe considerar a *lat* como la latitud y *lon* como la longitud; Δlat es la diferencia de; $lat2$ - $lat1$; $\Delta lon = lon2 - lon1$; R=6372.795477598 *km* que es el radio de la tierra.

El trabajo propuesto para generar el modelo de planeación escalable no debe ser planteado como una simple partición del grafo inicialmente determinado por $G=(V,E)$ y subdividido en conjuntos o subredes más pequeñas a través de una clusterización en la que se puede configurar una capacidad para cada centroide que entregaría las coordenadas de un PAD y la razón se debe a que los métodos de clusterización nos entregan un resultado diferente y con subconjunto desbalanceados en lo que se refiere al número de MIs por cada PAD.

La variable es determinante para el proceso de enrutamiento; por tal razón, se requiere encontrar la distancia más corta no solo entre cada MI sino también con cada PAD ya que en el despliegue de una red heterogénea los multi saltos no deben ser muy largos para evitar retardos en el envío de paquetes o mayor consumo de energía.

A continuación, se detalla el algoritmo utilizado en el modelo *OPDWHN-AMI:*

Algoritmo 1: *OPDWHN-AMI*

$Ns(j) = 0, \forall j \in \{1,2, \dots M\}, Cov = 0$
$Conjunto\ \mathbb{I}\ de\ Internet$
Input:
$Conjunto\ \mathbb{U}\ de\ DAPs, M = \| \mathbb{U} \|$
$Conjunto\ \mathbb{B}\ de\ EB, K = \| \mathbb{B} \|$
$Conjunto\ \mathbb{S}\ de\ MIs, N = \| \mathbb{S} \|$
$Parametro\ C_m\ capacidad\ de\ UDAPs$
***Step* 1**: $Construir\ matriz\ de\ costos\ G$
***Step* 2**: $for\ \mathbb{U}j, j \in \{1, \dots M\}$
$\Psi_i = calculando\ árbol\ (\mathbb{U}j, \mathbb{S}, C_m);$
***Step* 3**: $\mathbb{U}_r = SetCover(\Psi_1, \Psi_2, \dots \Psi_M)$
***Step* 4**: $[nextsalto, cost] = dijkstra(G, \mathbb{I})$
***Step* 5**: $for\ SM \in \mathbb{S}$
$nodo = SM$
$while\ nodo \in \mathbb{S}$
$nodo = nextsalto(nodo)$
$end\ while$
$Ns(nodo) = Ns(nodo) + 1$
$if\ nodo \in \mathbb{U}$
$Cov++$
$end\ if$
$end\ for$
***Step* 6**: $end\ for$
***Step* 7**: $retorna\ Cov/\| \mathbb{S} \|$
***Step* 8**: $retorna\ avg(Ns)$

La función Ψ implementada como un método recursivo de la heurística OPDWHN-AMI se presenta a continuación:

Algoritmo 2: Ψ – *Generación del árbol*

$function\ \Psi = calcular\ árbol\ (\mathbb{U}j, \mathbb{S}, C_m);$
***Step* 1**: $initializar\ Nsalto = 1, \Psi = \{\mathbb{U}j\}$
$nocon = \mathbb{S}$
$Ns(\mathbb{U}j) = \emptyset$
***Step* 2**: $while\ \| \Psi \| < C_m\ \&\&\ Nsaltos < Nsmax$
$dmin = \infty$
$for\ i \in \Psi$
$for\ j \in nocon$
$d = dist(i, j)$
$if\ d < dmin\ \&\&\ Ns(i) + 1 \leq Nsaltos$
$a = i, b = j, dmin = d$
$end\ if$
$end\ for$
$end\ for$
$if\ dmin < inf$
$\Psi = \Psi \cup \{b\}$
$nocon = nocon \backslash \{b\}$
$Ns(b) = Ns(a) + 1$
$else$
$Nsaltos = Nhop + 1$
$end\ if$
$end\ while$
***Step* 3**: $retorna\ \Psi$

El modelo de planeación de redes heterogéneas inalámbricas bajo las consideraciones anteriormente descritas define un problema combinatorio del tipo NP-Complete; por lo tanto, requiere un proceso heurístico para obtener un resultado cercano al óptimo que permita en la línea del tiempo seguir incluyendo nuevos MIs y PADs de ser el caso.

Mediante la tabla 2 se hace un resumen de las variables empleadas en el modelo de redes inalámbricas heterogéneas.

B. Red eléctrica soterrada

El presente modelo para planeación de redes de distribución eléctrica soterrada considera similar principio al aplicado para medición inteligente, con una variante en relación a la restricción para no permitir enlaces de cobertura radiales, sino que, permitir enlaces de cobertura lineales, evitando así, diagonales dentro del $G=(V,E)$ y, en consecuencia, se logra construir enlaces controlados sin cruces por los espacios delimitados por calles en todos sus lados, de esta manera, el despliegue de la red soterrada deberá seguir únicamente el camino preestablecido que refiere a calles y avenidas, en este caso, el problema advierte la necesidad ubicar de manera óptima el número de transformadores de distribución eléctrica; así como, el número de pozos de revisión que se requieran para el despliegue; para lo cual, se tomará información bajo normativas y estándares específicos para redes soterradas.

Para resolver el problema combinatorio definido como NP-Complete, planteado en este trabajo, se incluye un modelo heurístico que permite encontrar el costo mínimo para lograr el máximo porcentaje de cobertura de servicio para los usuarios finales. A continuación, en la ecuación 2 se presenta el costo total de cada transformador perteneciente a cada clúster, donde, es el número de elementos de cada clúster y C_p es la demanda individual de cada usuario dependiendo de la demanda estratificada tipo al que pertenezca. En la ecuación 3, C_{sub} es el costo total de la subestación, donde, es el número de clústeres existentes en el escenario. Un aspecto de mucha importancia es que, en cada iteración para obtener el costo de la subestación, se obtiene el costo acumulado por el algoritmo en cada transformador de distribución a desplegarse en el área de interés. De aquí en adelante, entiéndase por costo a los valores en MVAs o

kVAs (transformadores, subestaciones) y longitudes (conductores), según

$$C_{tr} = \sum_{k1=1}^{ze} C_p(k1)\ , \quad \forall\, ze, C_p \in \mathbb{R}^+ > 0; \qquad (2)$$

$$C_{sub} = \sum_{k1=1}^{ze}\sum_{p=1}^{le} C_{tr}\,(k1, p)\ , \quad \forall\, ze, le \in \mathbb{R}^+ > 0; \qquad (3)$$

En la ecuación 4, el costo total del conductor en la red de medio voltaje está representado por Cmv_{con}, la expresión d_a es la longitud del tramo de un punto i a un punto j, el factor de multiplicación k es una variable de costo por cada metro para la implementación de la red de medio voltaje y, por último, N es la longitud del vector que contiene las posiciones de los transformadores incluyendo la posición de la subestación. Por otro lado, la expresión N-1 representa el número de aristas mínimo para garantizar de conectividad a todos los transformadores desde la subestación.

$$Cmv_{con} = \sum_{a=1}^{N-1} d_a * k\ , \quad \forall\, N \in \mathbb{R}^+ > 1; \qquad (4)$$

En la ecuación 5, se presenta el costo total del conductor de la red de bajo voltaje multiplicado por el factor K explicado anteriormente. La expresión $d_{(h,a)}$ representa las distancias máximas requeridas en cada tramo para garantizar la conectividad desde la fuente hacia los nodos de servicio, la misma que, representa a la distancia más crítica, es decir, existirá distancias menores dentro de cada agrupación, pero no mayores.

$$Cbv_{con} = \sum_{h=1}^{le} \quad max \sum_{a=1}^{zise(ze)-1} d_{(h,a)} * k\ , \quad \forall\, d \in \mathbb{R}^+; \qquad (5)$$

En consecuencia, el problema de optimización, para planeación de redes eléctricas de distribución, se define de la siguiente manera mediante la ecuación 6:

$$\min \quad C_{tr} + C_{sub} + Cmv_{con} + Cbv_{con} \qquad (6)$$

Sujeto a:

$$\sum_{\phi_i \in \phi} \phi \leq \tau\,, \quad \forall\, \phi \in A(\beta); \forall\, \tau > 1 \tag{7}$$

$$\sum_{i=1}^{\Omega} \phi_i = \beta\,, \quad \forall i \in A(\beta); \forall \Omega \in \mathbb{R}^+ \tag{8}$$

$$d_{i,j} \leq \lambda, \quad \forall i,j \in A(\beta) \tag{9}$$

La ecuación 7, se expresa la restricción de capacidad, con lo cual, se limita el número de usuarios asociados a un transformador de distribución, donde, τ es la cantidad de elementos intra-clúster máximo permitido, $A(\beta)$ es el conjunto de usuarios emplazados en el área de interés y ϕ son los subconjuntos de agrupaciones. La ecuación 8, es una rutina de comprobación, en la cual, se verifica que la suma de los nodos de cada agrupación sea igual a la longitud de β de usuarios desplegados en el área. De esta manera, se garantiza que los subconjuntos creados cubran el 100% de los usuarios emplazados en el área de estudio. Con la ecuación 9 se garantiza que, en cada agrupación, únicamente existan elementos que cumplan con la restricción de distancia máxima permitida para asociar a un usuario a un transformador de distribución.

Tabla 2. Variables usadas para modelo OPDWHN-AMI

Símbolo	Descripción	Significado
$\mathbb{I}$	Internet	Conjunto internet
$\mathbb{B}$	Estación base cellular	Conjunto EB
$\mathbb{U}$		Conjunto PADs
$\mathbb{S}$	Medidores Inteligentes	Conjunto de MIs
C_m	Capacidad	Capacidad PAD
G	Matriz de conectividad	Matriz de costo
dist	distancia	Distancia haversine entre recursos
Cov	Cobertura	Cobertura de MIs
Ψ	Arbl	Steiner tree
$\mathbb{U}_r$	Set cover	Set cover/ radio de cobertura
Ns	MIs conectados	Promedio de MIs conectados
Ψ_i	Conjunto de MIs →PAD	Direct | multi saltos
nocon	No conectado	MIs no conectados
$Nsaltos$	# saltos	Saltos
$Nsmax$	Max # de saltos	Max # Saltos
$dmin$	Min distancia	Distancia mínima
a, b	Variable i, j	Variable temporal
$\mathbb{U}j$	subconjunto	Subconjunto de U

En la tabla 3 se presenta un resumen de las variables empleadas en el modelo de planeación de redes eléctricas de distribución.

Tabla 3. Variables usadas en modelo DORED

Símbolo	Descripción	Significado
C_{tr}	Costo	Costo de transformador de distribución
C_{sub}	Costo	Costo total de la subestación
Cmv_{con}	Costo	Costo total del conductor MV
Cbv_{con}	Costo	Costo total del conductor de la red de bajo voltaje
d	distancia	Distancias máximas requeridas en cada tramo
$dmax$	Distancia	Distancia máxima permitida
ze	Variable	Número de elementos de cada clúster
C_p	Demanda en kVA	Demanda individual según su estrato
le	Variable	Número de clústeres
d_a	Distancia	Longitud de un tramo de un punto i a un punto j
N	Variable	Contiene el valor de la longitud de del vector de Usuarios y transformadores
k	Variable costo	Costo por cada metro instalado en red MV
ϕ_i	Subconjuntos	Subconjuntos de agrupaciones
τ	Elementos de un clúster	Max # permitido en un clúster
β_i, β_j	Coordenadas	Coordenadas en latitud y longitud de usuarios
$A(\beta)$	Conjunto	Usuarios emplazados en área georreferenciada
Ω	Variable	Número de elementos a desplegarse
μ	variable	Longitud de la zanja
π	conjunto	Número de pozos a desplegarse

En la Figura 3 se presenta el flujograma del algoritmo Despliegue Óptimo de Redes Eléctricas de Distribución (DORED), el cual, contiene la metodología del modelo propuesto en el presente documento. Se parte del análisis y extracción de los datos georreferenciados obtenidos de la plataforma virtual gratuita OpenStreetMap. Posteriormente, mediante el software Matlab, se procede a preparar el conjunto de datos a ser utilizados, es decir, no toda la información es válida, puesto

que, en un archivo osm existen miles de coordenadas georreferenciadas (Latitud y longitud), de tal manera que, si no se hace una limpieza de datos, estos hacen que, el cpu-time que el ordenador emplee para encontrar la solución, aumente de manera significativa arrojando resultados poco confiables. Por lo tanto, es de mucha importancia ordenar la información y desechar datos que se encuentran fuera del área de interés para el despliegue óptimo de redes eléctricas de distribución.

Una vez que la información se encuentra clasificada se definen las variables a ser empleadas en el modelo. Estas variables son: coordenadas de viviendas, ubicación de subestaciones, sitios candidatos para emplazar los transformadores de distribución y datos de las vías disponibles para la construcción del mínimo árbol de expansión de la zanja, logrando generar con ello, la topología necesaria requerida en los primarios de los transformadores. En consecuencia, una vez definida las variables a emplear en el modelo, se genera una matriz de distancias de dimensión nxn (matriz cuadrada) que arma un mesh relacionando a todos los nodos entre sí. Para el cálculo de las distancias se emplea la ecuación de haversine, anteriormente mencionada, en la ecuación 1.

Una vez que se conoce la matriz de distancias se procede a generar varios escenarios posibles, es decir, se generan subconjuntos de soluciones previas. Además, estas soluciones previas se apoyan en técnicas de clusterización, específicamente en una variante del algoritmo de PRIM, el cual, construye árboles de mínima expansión. El aporte en la variante del algoritmo de PRIM, antes mencionada, radica en la integración de rutinas de verificación, el cual, permite ingresar restricciones de capacidad y rangos máximos permitidos de cobertura. Por lo tanto, mediante la variante de PRIM el modelo es capaz de agrupar nodos (usuarios finales, transformadores, subestaciones, sitios candidatos., etc.) y generar escenarios previos. El número de escenarios previos generados dependerá el número de sitios candidatos para emplazar transformadores de distribución, estos sitios candidatos son los pozos distribuidos a lo largo de la zanja, por la cual, pasará la red de medio voltaje.

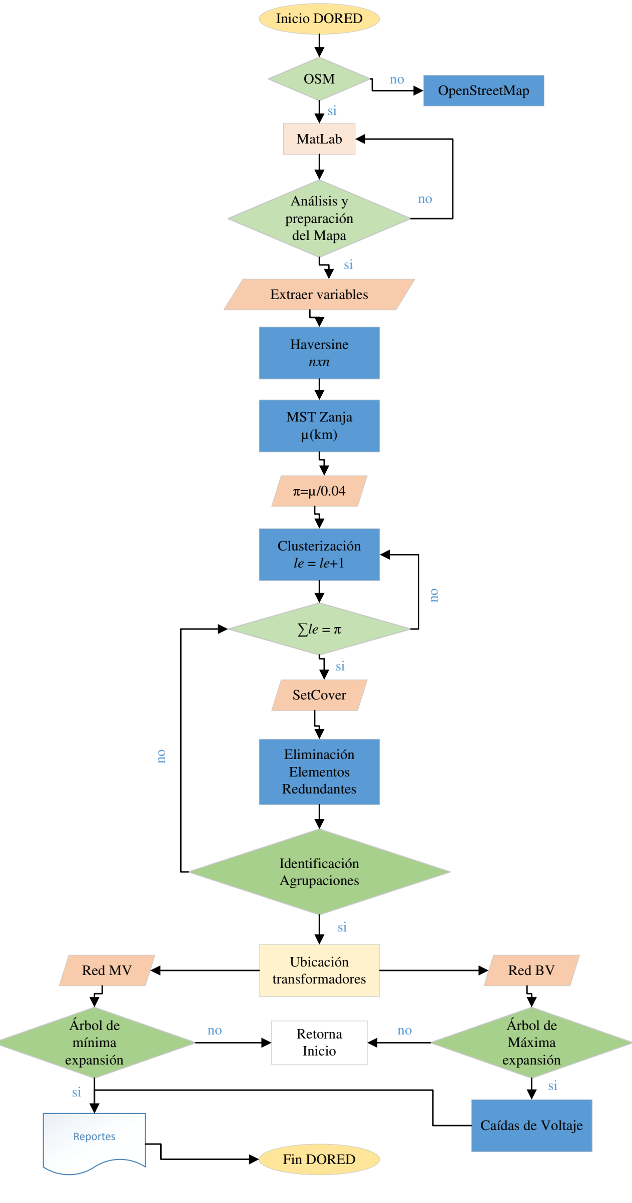


Figura 3. Flujograma del algoritmo dimensionamiento óptimo de redes eléctricas de distribución (DORED)

Un detalle de mucho interés, es que, la zanja podrá ser usada para el despliegue de fibra óptica. La distribución de los pozos se da a 40 metros y su emplazamiento se origina desde la subestación hacia toda la ruta que sigue la zanja en el caso de redes eléctricas soterradas. A manera de ilustración, si la zanja dispone de 5 km lineales desde la subestación hacia todos sus ramales el número de pozos requerido mínimo en el despliegue será de 125 excavaciones. Por lo tanto, los escenarios estarán representados por una matriz de *mnx* elementos, es decir, sus filas y columnas son de diferente longitud, donde, *n* representa a la cantidad de pozos y *m* el número de usuarios finales emplazados en el área de interés.

Una vez identificados los diferentes escenarios previos generados, se procede a ingresar dicha matriz al algoritmo de setcover, el cual, minimiza el número de escenarios generados garantizando la máxima cobertura a los usuarios finales. Es decir, se toman los subconjuntos que contengan el mayor número de nodos cubiertos hasta que todos los usuarios desplegados en el escenario queden asociados a algún subconjunto. En este punto se genera un inconveniente, recordemos que los escenarios son generados mediante una técnica de clusterización observando los mismos criterios de capacidad y cobertura en cada nueva iteración, lo que hace que, existan subconjuntos seleccionados como solución por el algoritmo del setcover que contengan los mismos elementos que otros subconjuntos, es decir, en la solución del algoritmo del setcover tendremos subconjuntos con elementos redundantes, lo que, en redes de distribución desemboca en la posibilidad de que un usuario pueda estar cubierto por más de 1 transformador de distribución, lo cual, en la práctica no es correcto.

Por lo tanto, una vez identificado los conjuntos se procede a eliminar los elementos redundantes siguiendo la metodología como se detalla a continuación; se ordenan los subconjuntos de mayor a menor y en cada iteración verificamos que el subconjunto siguiente no contengan los elementos del subconjunto predecesor de la matriz solución obtenida por el algoritmo Setcover. Este mismo criterio se aplica hasta hacer un barrido en toda la matriz, logrando con ello, eliminar los elementos redundantes garantizando que la capacidad de albergar usuarios sea la máxima posible en cada subconjunto.

Por lo tanto, como resultado tendremos el mínimo número de subconjuntos con la máxima cobertura posible a cada usuario final. Si un usuario se queda sin conexión a algún secundario de los transformadores a emplazarse, el diseñador conocerá toda la información necesaria, de tal manera que, pueda tomar decisiones adecuadas para otorgarle conectividad a algún secundario de los transformadores de distribución distribuidos en el área de interés. Una vez identificados los subconjuntos, con sus elementos redundantes eliminados, se procede a emplazar los transformadores de distribución. En consecuencia, el número de transformadores requeridos en el despliegue corresponde a la cantidad de subconjuntos resultantes del algoritmo del setcover.

Dicho de otra manera, tomando el mismo ejemplo anterior, en el que, se necesitaba 125 pozos para una longitud de 5 km de zanja, el algoritmo setcover arrojará como resultados posiblemente 80, 75, 50 (subconjuntos) unidades de transformadores requeridos en el escenario, esta variación dependerá exclusivamente de las restricciones de capacidad y cobertura ingresadas como restricciones en la generación inicial de los subconjuntos previos, antes de pasar por el algoritmo del setcover.

En consecuencia, si aumentamos la capacidad de agrupar elementos disminuiremos la cantidad de transformadores de distribución requeridos, pero, indudablemente aumentará la capacidad en kVAs requerida a instalar en dicho punto, caso contrario si, disminuimos la capacidad de agrupar aumenta el número de transformadores necesarios en el despliegue y disminuye la capacidad en kVAs de los transformadores necesarios en el despliegue.

Por lo tanto, se logra optimizar el número de transformadores mediante una técnica de clusterización, con la cual, se genera escenarios y posteriormente, mediante el algoritmo del setcover reducimos al máximo los subconjuntos necesarios para garantizar la máxima cobertura a los usuarios finales y, posteriormente se procede a ubicar los transformadores, en cada subconjunto, aproximadamente en el centro de masa, con la finalidad de garantizar distancias similares desde el transformador hacia sus puntos de carga más alejados.

En este punto, el que se ha logrado tener un conocimiento sobre de la distribución de los elementos de potencia en redes eléc-

tricas de distribución, se procede a generar árboles de mínima y máxima expansión a dos capas. Una capa en medio voltaje (mínima expansión) y la otra capa en bajo voltaje (máxima expansión), de tal manera que, la red de distribución óptima queda identificada. Por lo tanto, hasta este punto se logra proporcionar mediante el software Matlab y con la base de datos de OpenStreetMap un mapa de ruta, donde, mediante restricciones, tales como capacidad y cobertura, se garantiza que su costo sea el mínimo.

Una vez que se conoce la topología de la red eléctrica de distribución y cada uno de sus subconjuntos se procede a verificar el comportamiento de la caída de tensión desde el transformador hacia los nodos de carga más lejanos en cada subconjunto. Para ello, se ha otorgado cargas randómicas normalizadas a cada usuario. Estas cargas proporcionadas aleatoriamente corresponden a valores tipo considerando la introducción de cocinas de inducción y calentamiento de agua. Además, se ha dividido el área en sectores, es decir, se proporcionará valores randómicos de cargas tipo identificando los estratos a los que pertenece cada usuario, los valores de consumo en kVAs de cada estrato, se tomó de la Guía de diseño de la EEQ. Por lo tanto, el modelo es capaz de hacer el análisis de caídas de tensión desde los secundarios del transformador hacia sus puntos de carga más alejados y, a su vez, puede determinar la capacidad del transformador con valores ya estandarizados. El modelo nos permitirá ver y analizar el comportamiento, si los usuarios aumentan su demanda, ayudando con eso a planear posibles expansiones por incremento de energía demandada, que con el pasar del tiempo, su tendencia está en aumento. Finalmente, el modelo entrega un reporte ejecutivo de los elementos requeridos para la construcción de la red eléctrica de distribución.

4. Análisis de resultados

A. Red heterogénea inalámbrica para medición inteligente

en esta sección, se demuestra como el modelo de planeación de una red heterogénea inalámbrica que despliega PADs y los MIs que finalmente se articular hasta lograr la conectividad con una estación base celular. Se evalúan parámetros que dan cuenta del

rendimiento en términos de minimización de los recursos. La Figura 4 muestra el despliegue de los PADs candidatos, los mismos que se ubican en las intersecciones de calles y avenidas, lugar donde se ubican postes de alumbrado público y donde finalmente se podrán instalar los PADs requeridos para dar cobertura a los MIs de cada una de las viviendas, considerando la capacidad de los PADs en función de la tecnología inalámbrica en cada MI o PAD.

El escenario advierte un número inicial de PADs con =57, un número de MIs =349 y estación base celular =1. El escenario se ha dividido por un proceso de teselación con Delaunay y Voronoi para segmentar las áreas que pueden determinar los estratos o grupo de viviendas, según detalla la Figura 5

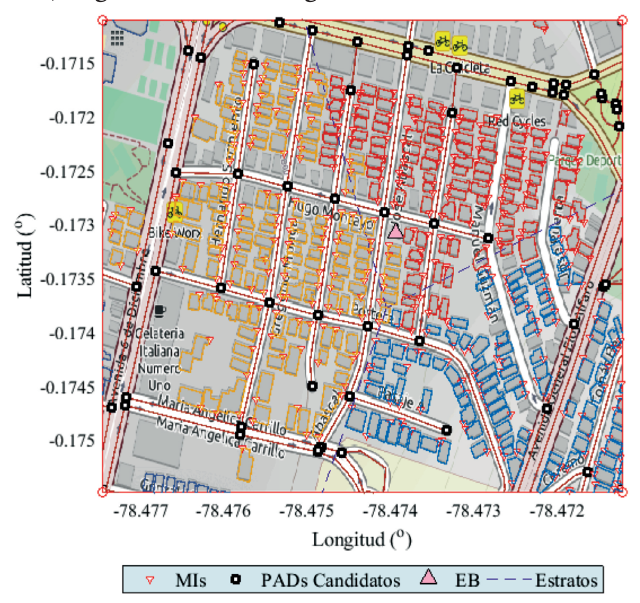


Figura 4. Escenario de planeación de una red heterogénea inalámbrica

Además, la Figura 5 indica la reducción del número *M* de PADs que se requieren para cubrir un número *N* de MIs; se aprecia también el proceso de teselación realizada con Delaunay y Voronoi para dividir en 3 áreas o estratos que dividen los diferentes tipos de clientes que puede segmentar una empresa de distribución eléctrica; en este caso, cada estrato toma un color diferente.

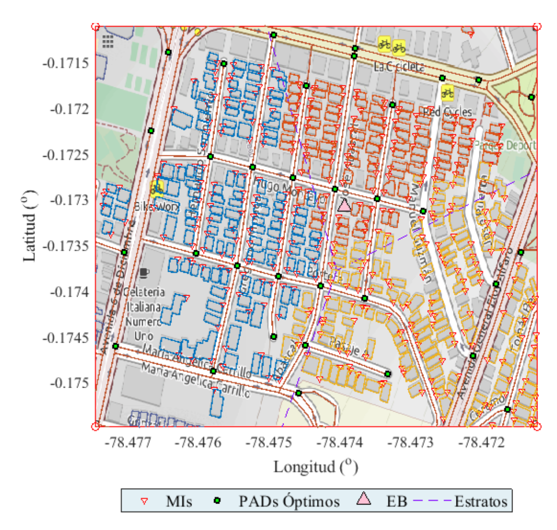


Figura 5. Optimización de recursos – minimización de PADs

La Figura 6 representa la malla factible lograda para el escenario propuesto con una *dmin*=49*mts*; de esta manera, se consigue la conectividad del grafo $G=(V,E)$. Con esta visión previa de los posibles enlaces inalámbricos se procede a buscar el mínimo árbol de expansión que se podría lograr.

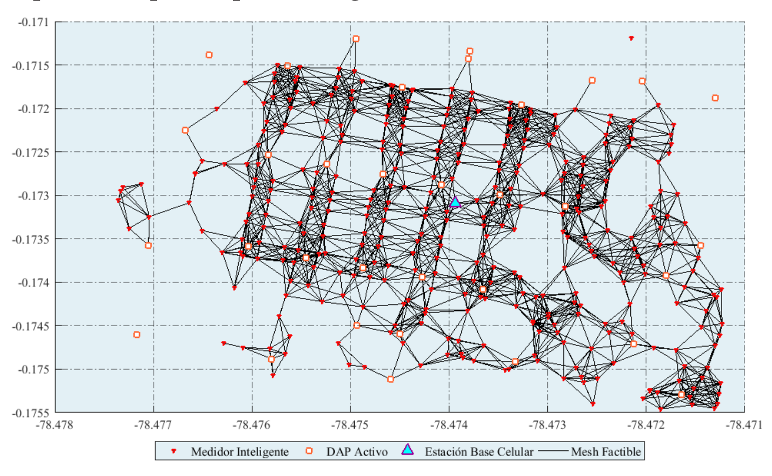


Figura 6. Malla factible de conexiones posibles con restricción de distancia

La Figura 7 muestra la reordenación inversa Cuthill-McKee para mover todos los elementos distintos de cero más cerca de la diagonal, reduciendo el "ancho de banda" de la matriz original G.

La Figura 8 muestra el mínimo árbol de expansión, luego de haber resuelto el problema del setcover que logra la mayor cobertura; es decir, maximiza la cobertura de todos los MIs con la mínima solución encontrada al generar el mínimo árbol de expansión a través del algoritmo Ψ para generación del árbol.

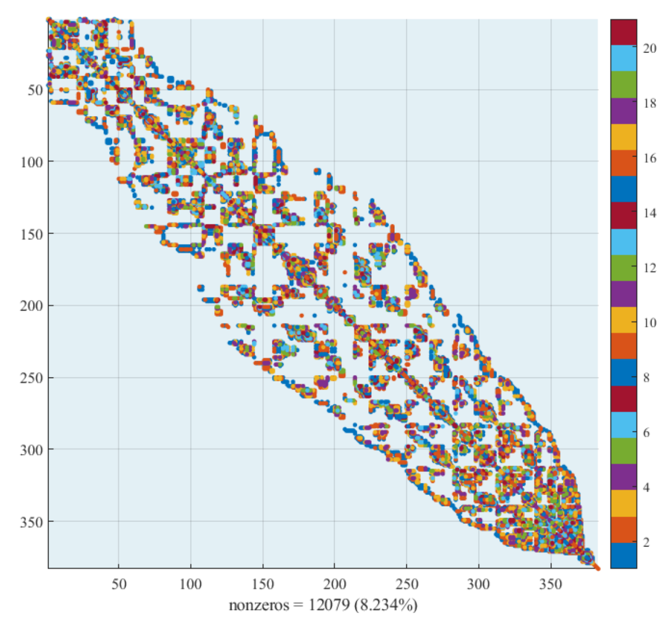


Figura 7. Reordenación inversa Cuthill-McKee

La Figura 9 representa los subconjuntos obtenidos al haber incluido tres tipos de tecnologías inalámbricas y en conjunto se puede visualizar los recursos que son agrupados y que a través de un PAD finalmente llegan hasta una estación base celular, la representación de este proceso se encuentra en color celeste. Los enlaces de tecnología inalámbrica WiFi se encuentran en color rosado; los enlaces entre PAD en color verde, para este caso la tecnología inalámbrica puede ser WiFi o Celular dependiente de las características de la zona y los recursos existentes en términos de cobertura celular.

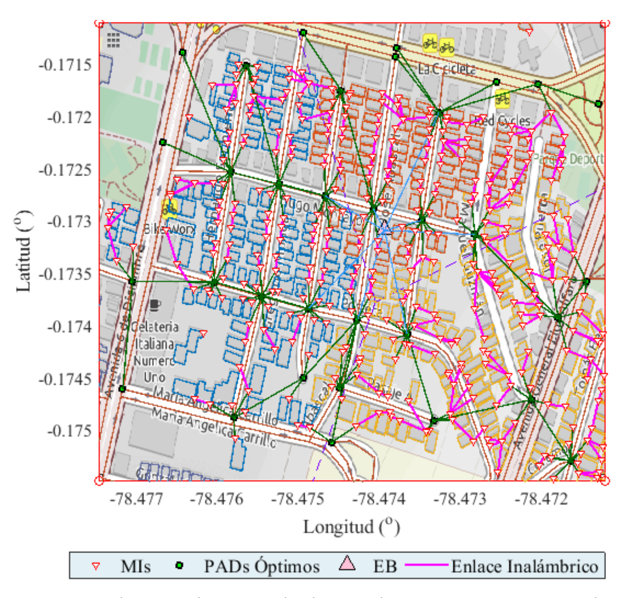


Figura 8. Red Despliegue de la Red Heterogénea Inalámbrica

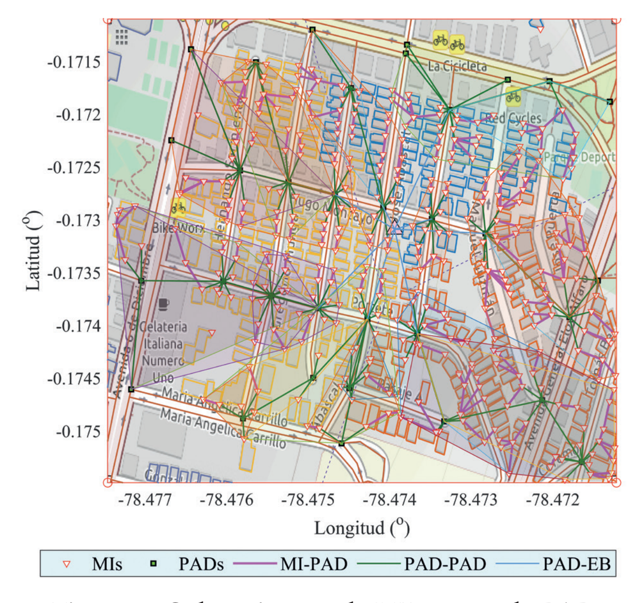


Figura 9. Subconjuntos de MIs por cada PAD

B. Red de Distribución Eléctrica Soterrada

En la Figura 10 se presenta la disposición del escenario, el cual, se empleará para desplegar transformadores eléctricos de distribución. El escenario propuesto dispone de 3 estratos, ilustrados en la Figura 10 mediante tres colores diferentes. El criterio de estratificación se lo llevó a cabo mediante el algoritmo de K-medoids tomando en consideración la guía de diseño de redes de distribución de la Empresa Eléctrica Quito (EEQ) para otorgar demandas máximas tipo por cada usuario. Los estratos empleados en el presente documento son: E, D, C y B. Es de mucha importancia mencionar que, el modelo, permite modificar dichos parámetros para poder ajustarse a diferentes realidades o disposiciones topográficas. A cada usuario de cada estrato se le otorgó un consumo referente máximo randómico normalizado.

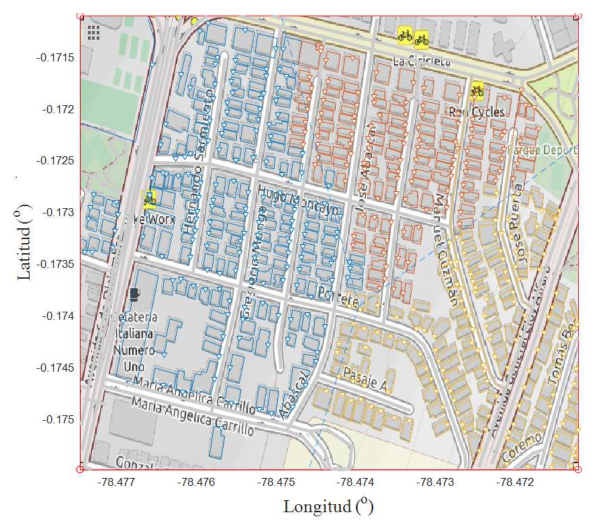


Figura 10. Escenario inicial considerando estratificación

En la tabla 4 se ilustra un cuadro resumen de los resultados obtenidos del despliegue óptimo de transformadores para satisfacer de servicio eléctrico a los abonados al sistema de distribución, considerando, diferentes escenarios. La heurística propuesta, para resolver el problema combinatorio de ubicación de transformadores observando restricciones de capacidad y cobertura, se valida mediante la tabla

4, donde, se puede observar los diferentes escenarios resueltos por el modelo. Se puede apreciar además que, a medida que, la capacidad aumenta los parámetros de las columnas 2, 4 y 5 también lo hacen. En el caso de la columna 2 de la tabla 4 se desvela que, al incrementar la capacidad la variación de la cobertura no sigue ningún patrón, es decir, no es ni directamente proporcional ni inversamente proporcional a la capacidad, por lo tanto, el porcentaje de cobertura no depende de la capacidad de agrupamiento del transformador, sino que, depende de la disposición geográfica de los usuarios finales, puesto que, la ubicación involucra directamente a la variable distancia, que es, una restricción tonada en cuenta en el presente modelo.

Tabla 4. Planeación y despliegues obtenidos variando capacidad de agrupamiento de cada transformador

Capacidad Máxima (unidades)	Cobertura (%)	Cap. Subestación (MVAs)	Transformadores Requeridos (kVA/cantidad)	kVA-m máximo
10	83.77	1.42	45/19	493.17
			60/5	325.43
			15/7	103.34
			50/5	400.12
			30/5.	476.17
15	82.56	1.44	100/7	513.22
			75/5	448.40
			60/7	603.94
			45/2	242.16
			30/2	219.40
			50/1	212.47
			15/4	222.31
18	90.79	1.42	100/9	636.45
			75/6	592.41
			60/1	694.19
			45/2	458.53
			30/2	222.26
			15/7	171.51
			112.5/2	596.00

Capacidad Máxima (unidades)	Cobertura (%)	Cap. Subestación (MVAs)	Transformadores Requeridos (kVA/cantidad)	kVA-m máximo
20	92.49	1.44	100/10 112.5/ 3 125/ 2 60/ 1 75/ 2 45/ 1 15/ 6	815.57 641.61 653.40 248.59 1052.02 347.48 225.66
25	94.18	1.42	112.5/5 125/2 150/3 100/3 75/1 45/2 50/1 15/6	525.35 555.68 674.59 1114.74 489.63 244.13 186.96 219.51
30	97.57	1.43	150/5 160/2 112.5/3 125/2 75/1 100/1 60/1 15/6	1219.87 747.23 679.50 726.31 383.93 467.79 253.82 231.20
35	92.25	1.43	160/2 200/3 150/5 112.5/1 15/6	845.92 893.25 696.50 388.98 223.93
43	83.7	1.44	200/5 250/2 15/6 45/2 30/1	1190.86 1026.78 178.27 480.89 222.32

En la columna 3, de la tabla 4, se presenta la capacidad calculada por el algoritmo de la subestación requerida en MVAs. En párrafos anteriores se ha explicado que, a cada usuario, se otorgó un consumo máximo tipo randómico, lo que hace que, en cada iteración la heurística obtenga diferentes resultados muy próximos entre sí. Dichos resultados son controlados, es decir, se otorga por cada estrado un nivel mínimo y máximo de consumo randómico a cada usuario. Por lo tanto, la heurística propuesta no coloca cargas rígidas o únicas a los usuarios finales y además permite identificar, por la posición geográfica, el estrato al que pertenece cada abonado a la red eléctrica de distribución. En consecuencia, la capacidad de la subestación no depende de la capacidad en MVAs estandarizado de cada transformador, sino que, depende de la carga total instalada y del número y tipo de usuarios. Por lo tanto, el valor obtenido de la tabla 4 en la columna 3, debe ser estandarizado de acuerdo a la disposición del mercado en MVAs de subestaciones de distribución.

En la columna 4 de la tabla 4 se aprecia que, a medida que la capacidad aumenta, disminuye la cantidad de transformadores y aumenta la necesidad de incrementar la capacidad en MVAs requerida para satisfacer de servicio a usuarios finales asociados a la red eléctrica de distribución. Por lo tanto, queda demostrada la funcionalidad y validez de la heurística propuesta. En la columna 5 se presenta la métrica en kVA-m máximo desde un transformador hasta su punto más lejano, es decir, se presenta el peor de los casos. Mediante la métrica, expuesta anteriormente, se puede dimensionar el calibre del conductor a emplearse, considerando la caída de voltaje en los extremos más alejados desde la fuente (transformador de servicio), en bajo voltaje, mediante la relación kVAs-m calculados/kVAs-m soportados.

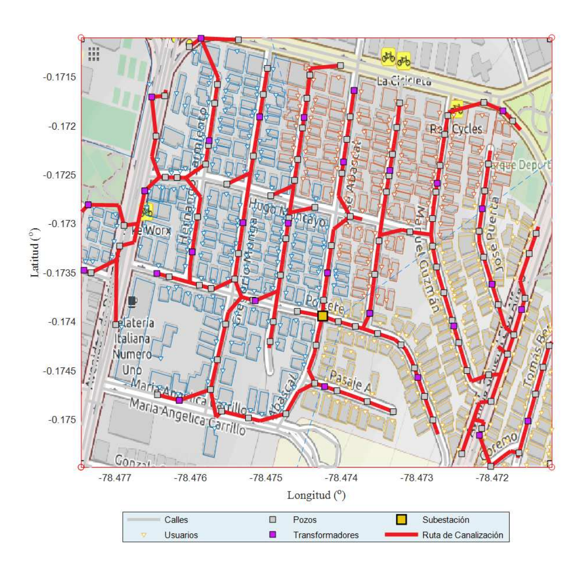


Figura 11. Planeación óptima de una red de distribución soterrada de energía eléctrica considerando capacidad y cobertura

En la Figura 11 se presenta el resultado de la planeación obtenida por la heurística propuesta en el presente documento. Mediante la figura, anteriormente mencionada, se puede apreciar el mapa de ruta a emplearse garantizando el mínimo costo. El escenario de la Figura 11. dispone de N=349 usuarios distribuidos geográficamente y corresponde a los resultados obtenidos en la tabla 4 con capacidad máxima de albergar a 18 usuarios. Además, la heurística planteada proporciona datos de variables que son determinantes en el momento de elegir sobre la solución a ejecutarse en el proceso de planeación, tales como: número de pozos requerido (115), longitud de canalización o zanja (4.614 km) y demanda total máxima de cada estrato: 1 (431.44 kVA), 2 (371.22 kVA) y 3 (626.04 kVAs). En consecuencia, el modelo propuesto es capaz de modelar de manera dinámica el despliegue de transformadores eléctricos dentro de un área georreferenciada proporcionándonos parámetros valiosos de diseño con los cuales, el ingeniero eléctrico, puede tomar decisiones observando parámetros técnicos. Por lo tanto, en el presente documento se desvela una herramienta muy prometedora de diseño en

redes eléctricas de distribución soterrada que ayude al diseñador a tomar decisiones, no solo más certeras, sino que, garantice la continuidad y calidad del servicio eléctrico a los usuarios finales pertenecientes a la red eléctrica de distribución.

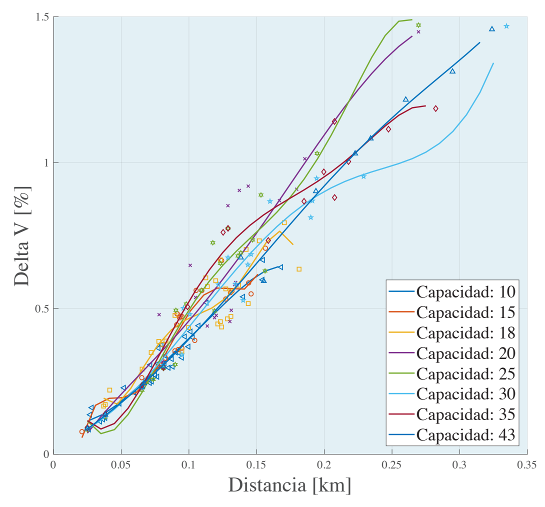


Figura 12. Máximas caídas de voltaje en la red secundaria en diferentes escenarios considerando capacidad y rangos de cobertura

En la Figura 12 se aprecia el comportamiento del voltaje en función de la distancia en la red de medio voltaje. Es evidente que la variación de voltaje, expresada en porcentaje, aumenta a medida que la distancia también lo hace, esto se debe a que, a mayor longitud del conductor existe mayores pérdidas por efecto Joule. Los datos de la Figura 12, en el escenario ensayado, están distribuidos de manera dispersa, puesto que, como se explicó en párrafos anteriores, los datos obtenidos no siguen ningún patrón a medida que incrementa la capacidad, sino que, depende del escenario, la capacidad de agrupamiento observando las restricciones propias del modelo propuesto y la capacidad de conducción del conductor a emplear. Además, la variación de voltaje no solo depende de la distancia, sino que, también depende de las distancias acumuladas a lo largo del tramo hasta llegar al usuario más alejado del circuito, del lado secundario del transformador al que pertenece. En consecuencia, se

puede apreciar que a medida que se aumenta la capacidad los porcentajes de variación de voltaje también lo hacen y esto sucede, ya que, al aumentar la capacidad de albergar usuarios por un transformador de manera conjunta aumenta las distancias desde un transformador hacia su nodo más lejano de servicio. Por último, mediante la Figura 12, se aprecia que, a cada transformador en sus diferentes escenarios con sus respectivas variaciones de voltaje generales, de lo que, se puede inferir que ningún transformador supera los límites máximos de caídas de voltaje desde la fuente hacia el usuario más alejado del transformador en cada escenario propuesto.

En la Figura 13 se ilustra el modelado de la subestación empleada en el proceso de simulación de la red eléctrica en los primarios de los transformadores otorgándole criterios de confiabilidad, lo cual, naturalmente eleva los costos de implementación. En el modelado del mapa de ruta obtenido por la heurística mediante el software MATLAB, se ha extraído la información necesaria como: cargas concentradas en cada transformador, número y capacidades requeridas en MVAs de los transformadores para satisfacer a la demanda en cada tramo, dimensionamiento de la subestación requerida, número de alimentadores y longitudes en cada tramo, ya sea, en medio y bajo voltaje. En consecuencia, una vez obtenidos las variables necesarias para el diseño de despliegue óptimo de transformadores considerando las demandas totales tipo por cada abonado, mediante el software Cymdist se procede a simular un flujo de potencia. Mediante la simulación del flujo de potencia podemos verificar el estado de la red eléctrica en medio voltaje con las condiciones proporcionadas en el mapa de ruta obtenido de la heurística propuesta. Por lo tanto, en la Figura 13. se verifica, además de la configuración de la subestación, el flujo de potencia, logrando con ello, obtener los datos de las variables de voltajes en p.u. con sus respectivos ángulos, corrientes en los conductores, potencia activa y potencia reactiva, con lo que, se puede ejecutar procesos de planeación en redes eléctricas de distribución. Es muy importante mencionar que, los valores en p.u. son adimensionales y pueden representar a cualquier magnitud física, ya sea, voltaje, corriente, impedancias etc., en consecuencia, cada valor en p.u. adimensional depende de su valor base real de cualquier parámetro eléctrico al que nos refiramos, para convertir un valor en p.u., en valores reales.

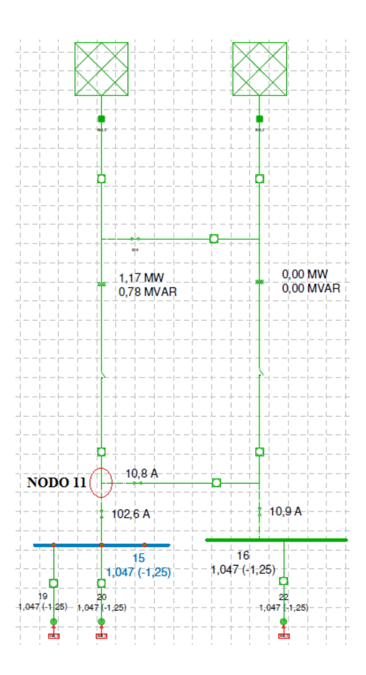


Figura 13. Modelado de la subestación mediante software cymdist



Figura 14. Diagrama unifilar trifásico en media tensión mediante software cymdist

Mediante la tabla 4 y la Figura 13. se valida el funcionamiento de la heurística propuesta en el presente trabajo, puesto que, mediante el triángulo de potencias, con los datos de la Figura 13, nos permite encontrar, en MVAs, la capacidad de la subestación requerida para satisfacer la demanda bajo las consideraciones descritas anteriormente. En consecuencia, con los datos de potencia activa y reactiva obtenidos de la simulación en el software Cymdist, calculamos el valor de la potencia aparente, cuyo valor es 1.41 MVAs, mientras que, si nos percatamos en la tabla 4 en la fila con capacidad de 18 usuarios, la capacidad calculada por el modelo en MVAs es de 1.42. Por lo tanto, los valores obtenidos en el software Cymdist y los calculados mediante el modelo propuesto en el software MATLAB son semejantes, de tal manera que, se logra demostrar y validar la funcionalidad del modelo propuesto garantizando altos niveles de confiabilidad en los resultados obtenidos.

En la Figura 14 se muestra el diagrama unifilar obtenido en la simulación del flujo de potencia mediante el software Cymdist. En consecuencia, el mapa de ruta georreferenciado obtenido del software Matlab es posible simular, dicho mapa de ruta, en el software Cymdist con la finalidad de verificar el comportamiento de la red eléctrica en régimen permanente. Por lo tanto, una vez obtenido el mapa de ruta, mediante el software Cymdist, se puede realizar todo tipo de análisis pertinente en un proceso de planeación de redes de distribución, ya sea, soterrado, aéreo o mixto. Por lo tanto, la importancia de la propuesta planteada como una solución para despliegue de redes de distribución del presente documento, radica en que, mediante la observación de restricciones emplace los transformadores de potencia en los sitios idóneos garantizando minimizar costos maximizando la calidad y seguridad energética proporcionada a los usuarios finales.

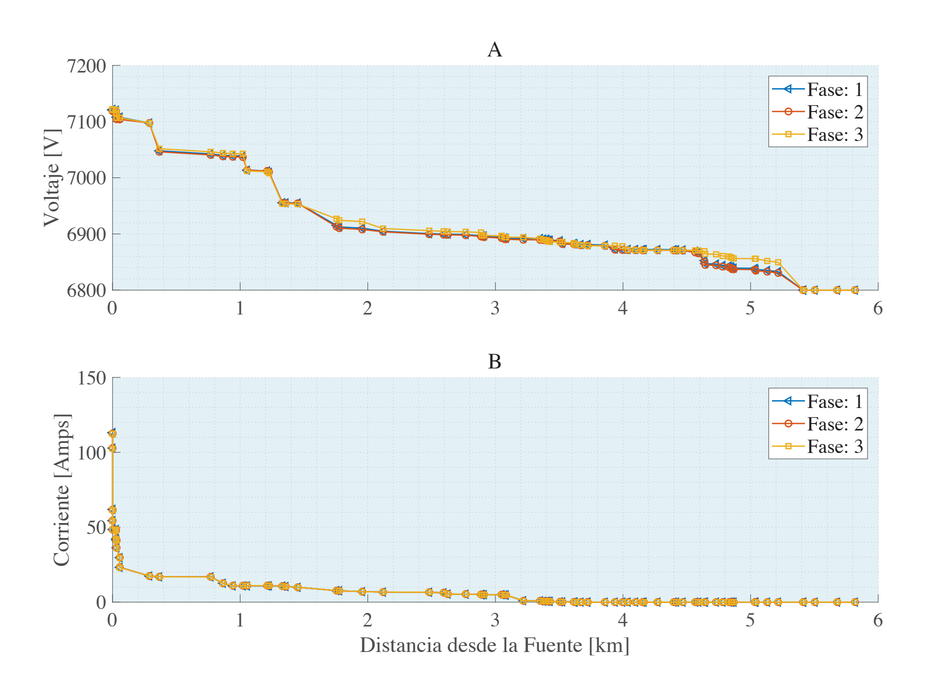


Figura 15. Perfil de voltaje y corriente circulantes en cada tramo desde la fuente

En las métricas de la Figura 15 se puede apreciar el comportamiento del voltaje y la corriente en función de la distancia. La longitud máxima desde la fuente hacia el transformador más alejado es aproximadamente 580 m. En la Figura 15A se expresa que a medida que un transformador se aleja de la fuente existe caídas de voltaje en cada tramo, además, las caídas de voltaje en cada son similares, lo que pone de manifiesto que, para la presente explicación se ha modelado un sistema trifásico con cargas repartidas equilibradas en cada fase. Por otro lado, la Figura 15B nos muestra las máximas corrientes soportadas en la fuente y a medida que la distancia aumenta esas corrientes van disminuyendo, esto sucede ya que, en la fuente se refleja el consumo total de la zona a la que, dicha fuente, sirve. Por lo tanto, a medida que la distancia de servicio aumenta la corriente disminuye, ya que, la corriente nominal desde la fuente, en cada nodo primario, dependiendo de la topología se va dividiendo.

En la Figura 16 A y B sucede algo similar que lo explicado en el párrafo anterior. En consecuencia, a medida que la distancia aumenta, tanto la potencia aparente como la potencia reactiva, disminuyen. Estas métricas nos permiten observar el estado de la red eléctrica de distribución y planificar la colocación de bancos de condensadores en caso de ser necesario.

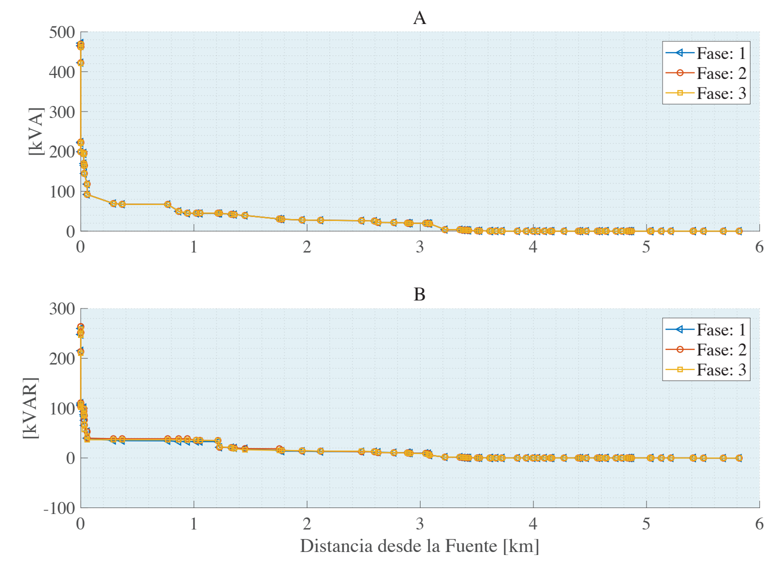


Figura 16. Perfiles de potencia aparente y reactiva desde la fuente en función de la distancia

En la Figura 17 se ilustra el diagrama unifilar al aplicar, mediante el software Cymdist, el módulo de análisis de fallas "Huecos de Tensión". Cada color representa la magnitud en p.u del hueco de tensión en cada tramo. En consecuencia, el color rojo simboliza mayor que 0 y menor que 0.1, el color naranja mayor que 0.2 y menor que 0.3, el color verde mayor que 0.5 y menor que 0.6 y por último el color azul mayor que 0.9 y menor que 1.



Figura 17. Análisis de huecos de tensión en la red de media tensión mediante software cymdist

Sabemos que, los huecos de tensión son perturbaciones en los sistemas eléctricos que afecta a la calidad del suministro de energía eléctrica. En consecuencia, el análisis de los huecos de tensión se produce en forma repentina en un punto particular de la red, para el caso, se ha hecho el análisis de hueco de tensión en el Nodo 11 (Ver Figura 13). Por lo tanto, los huecos de tensión se producen por la reducción súbita de la tensión recuperando sus valores habituales en un intervalo corto de tiempo y la causa habitual para que esto suceda son los cortocircuitos.

Por lo tanto, los huecos de tensión pueden influir en la impedancia y localización de falla, la forma de actuación de las protecciones, la conexión de la carga y el tipo de fallo. Además, provoca fallos en los equipos conectados a la red eléctrica, puesto que, existen dispositivos que necesitan cierta calidad de alimentación y pueden verse perjudicados en su rendimiento.

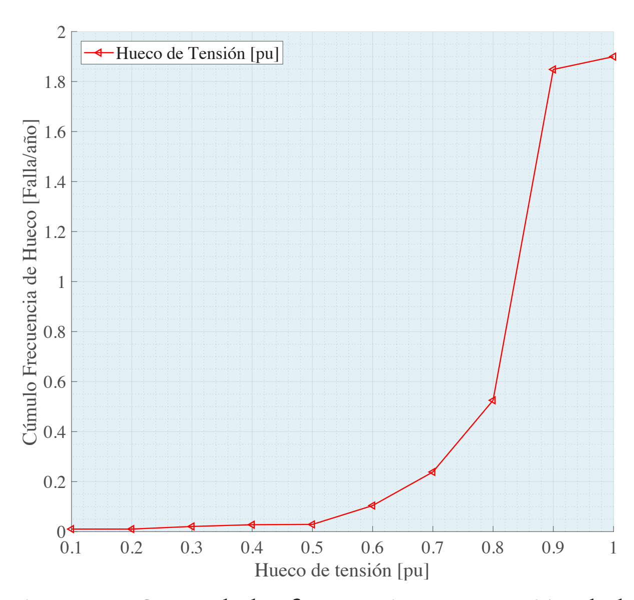


Figura 18. Suma de las frecuencias y su tensión de hueco

Como se ha explicado en párrafos anteriores un hueco de tensión está relacionado con la reducción del valor rms por un periodo corto de duración. Los valores relacionados en p.u. de la reducción de la componente fundamental de tensión son de 0.1 a 0.9, con duración mayor que medio ciclo y menor de 1 minuto.

En la Figura 18 se ilustra en el eje de las ordenadas la suma de huecos de tensión esperados por año y en el eje de las abscisas las magnitudes de los diferentes huecos de tensión. En consecuencia, el cúmulo de frecuencia de huecos de tensión da cuenta, de cuántos huecos de tensión de valores remanentes menores o iguales a cierto umbral ocurrieron en un periodo de tiempo.

Es necesario mencionar que, el cúmulo de frecuencia al estar relacionado al número de fallas durante un año (tasa de falla) Cymdist modela este fenómeno en la red de distribución eléctrica con el ingreso de datos de las tasas de falla tanto de las barras como porciones de las líneas, cada posición de falla puede tener una tasa de falla diferente, la misma que, se puede relacionar directamente con el tipo de instalación. Las tasas de fallas son también diferenciadas

en los tipos de falla de podrían producirse en los sistemas eléctricos de potencia, trifásico, monofásico, bifásico y bifásico a tierra.

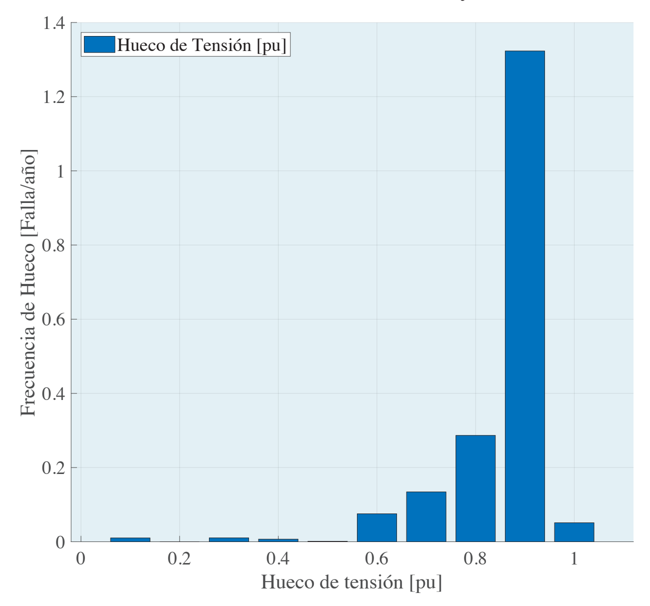


Figura 19. Número de fallas por año con respecto a la magnitud del hueco

De manera similar, en la Figura 19. se ilustra el número de fallas ocurridas al año con respecto a la magnitud del hueco de tensión. Es de mucha importancia citar que, el análisis de huecos de tensión simulado mediante el software Cymdist se aplicó en el Nodo 11, mismo que, se ilustra en la Figura 13.

Por lo tanto, queda demostrado que en el presente documento se logra, mediante el software Matlab, implementar una heurística capaz de dar solución óptima al problema de planeación de redes de distribución considerando restricciones de capacidad y cobertura otorgando confiabilidad a la solución obtenida, puesto que, se ha validado dicha solución mediante el software Cymdist. En consecuencia, la heurística propuesta es una herramienta de mucha utilidad para proporcionar soluciones optimas de planeamiento, utilizando teoría de grafos, en redes eléctricas de distribución. Además, mediante simuladores eléctricos, el diseñador, puede garantizar la viabilidad del proyecto minimizando los costos de implementación.

5. Conclusiones

El modelo de planeación de redes heterogéneas inalámbricas conseguido permite visualizar un mapa de ruta para que empresas de distribución eléctrica avizoren un panorama mucho más claro en lo que respecta al despliegue de la infraestructura de medición avanzada que dará paso a medición inteligente de energía eléctrica.

El modelo propuesto es escalable y flexible, lo que, permite incorporar nuevos medidores inteligentes; incorporar tecnologías inalámbricas según sea la necesidad; además se puede aplicar a zonas urbanas, suburbanas y rurales lo que convierte al modelo en una aplicación real para minimizar el tiempo en el despliegue de medición inteligente en una ciudad o país.

El trabajo propuesto advierte que no se requiere usar algoritmos de clusterización para agrupar a medidores inteligentes en un punto de agregación de datos debido a la aleatoriedad con la que este tipo de algoritmos trabaja, la metodología definida en un previo árbol de mínima expansión, un setcover y finalmente un enrutamiento usando el algoritmo Dijkstra en su versión desde la estación base celular a los medidores inteligentes mantiene un tiempo de ejecución adecuado para verificar resultados con diferente número de medidores inteligentes.

Con la presente investigación se ha validado el funcionamiento y la confiabilidad de la heurística propuesta en la solución óptima de despliegue de transformadores de potencia para redes de distribución en media y bajo voltaje observando restricciones de capacidad y cobertura. Es un modelo capaz de calcular caídas de voltaje en la red de baja tensión para cada transformador emplazado, garantiza la observancia de las restricciones, lo que hace que, el mapa de ruta obtenido por el modelo sea el óptimo para su respectiva implementación. Además, mediante teoría de grafos y rutas mínimas se garantiza que la topología o mapa de ruta obtenido por el presente modelo sea el mínimo y, por lo tanto, el costo de implementación, de igual manera, sea el mínimo. El modelo, no solo proporciona el mapa de ruta, sino que, proporciona longitud en kilómetros de la ruta que debe seguir el tendido eléctrico, cantidad de pozos o postes, cantidad y capacidad de los transformadores a

emplazarse con valores estandarizados, ubicación georreferenciada de los dispositivos dispersos en el área de interés, considera además estratificación en base a los consumos tipo y, además, proporciona la capacidad requerida de la subestación en MVAs. Por lo tanto, el modelo es capaz de dar solución al problema combinatorio de complejidad NP-Complete. Por último, mediante el software Cymdist validamos la funcionalidad del modelo y la capacidad de proporcionar topologías o mapas de ruta óptima para posteriormente mediante el simulador analizar el comportamiento estrictamente eléctrico considerando parámetros eléctricos no tomados en cuenta en la heurística propuesta, tales como, impedancias, coseno phi, configuraciones de subestaciones, etc.

En futuros trabajos implementaremos restricciones de caídas de voltaje para la selección de la ubicación georreferenciada de los transformadores de distribución y, además, se incluirá flujos de potencia en el algoritmo.

Agradecimientos

Este trabajo ha sido producido gracias al apoyo del GIREI - Grupo de Investigación en Redes Eléctricas Inteligentes de la Universidad Politécnica Salesiana Ecuador bajo el proyecto Planeación y Despliegue Óptimo de Redes de Distribución Eléctrica Aéreas y Soterradas.

Referencias

[1] E. Inga, S. Céspedes, R. Hincapié, and A. Cárdenas, "Scalable Route Map for Advanced Metering Infrastructure Based on Optimal Routing of Wireless Heterogeneous Networks," IEEE Wirel. Commun., no. April, pp. 1–8, 2017.

[2] M. Campaña, E. Inga, and R. Hincapié, "Optimal placement of universal data aggregation points for smart electric metering based on hybrid wireless," in CEUR Workshop Proceedings, 2017, vol. 1950.

[3] A. Peralta, E. Inga, and R. Hincapié, "Optimal Scalability of FiWi Networks Based on Multistage Stochastic Programming and Policies," J. Opt. Commun. Netw., vol. 9, no. 12, p. 1172, 2017.

[4] F. Aalamifar, G. N. Shirazi, M. Noori, and L. Lampe, "Cost-efficient data aggregation point placement for advanced metering infrastructure," in Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on, 2014, pp. 344–349.

[5] G. Rolim, D. Passos, I. Moraes, and C. Albuquerque, “Modelling the Data Aggregator Positioning Problem in Smart Grids,” 2015 IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput., pp. 632–639, 2015.
[6] D. Li, Z. Aung, J. R. Williams, and A. Sanchez, “Efficient Authentication Scheme for Data Aggregation in Smart Grid with Fault Tolerance and Fault Diagnosis,” pp. 1–8, 2011.
[7] J. Inga, E. Inga, R. Hincapié, and G. Cristina, “Optimal Planning for Deployment of FiWi Networks based on Hybrid Heuristic Process,” Lat. Am. Trans. IEEE (Revista IEEE Am. Lat., vol. 15, no. 9, 2017.
[8] A. Peralta, E. Inga, and R. Hincapié, “FiWi Network Planning for Smart Metering Based on Multistage Stochastic Programming,” Lat. Am. Trans. IEEE (Revista IEEE Am. Lat., vol. 13, no. 12, pp. 3838–3843, 2015.
[9] E. Inga and R. Hincapié, “Matched Channel Allocation for Advanced Metering Infrastructure based on Cognitive Mobile Virtual Network Operator,” Lat. Am. Trans. IEEE (Revista IEEE Am. Lat., vol. 14, no. 4, pp. 1780–1785, 2016.
[10] G. Rajalingham, Q.-D. Ho, and T. Le-Ngoc, “Evaluation of an efficient Smart Grid communication system at the neighbor area level,” in Consumer Communications and Networking Conference (CCNC), 2014 IEEE 11th, 2014, pp. 426–431.
[11] J. Inga, E. Inga, R. Hincapié, and C. Gómez, “Evaluación de la Infraestructura de Medición y la Respuesta de la Demanda,” Rev. Técnica Energía, no. 12, pp. 262–269, 2016.
[12] M. Maier, “FiWi access networks: Future research challenges and moonshot perspectives,” 2014, pp. 371–375.
[13] E. Inga-Ortega, A. Peralta-Sevilla, R. C. Hincapie, F. Amaya, and I. Tafur Monroy, “Optimal dimensioning of FiWi networks over advanced metering infrastructure for the smart grid,” 2015 IEEE PES Innov. Smart Grid Technol. Lat. Am. (ISGT LATAM), pp. 30–35, 2015.
[14] E. Inga-ortega, A. Peralta-Sevilla, R. Hincapié, F. Amaya, and I. Tafur, “Optimal Dimensioning of FiWi Networks over Advanced Metering Infrastructure for the Smart Grid,” Innov. Smart Grid Technol. Lat. Am. (ISGT LATAM), 2015 IEEE PES, pp. 30–35, 2015.
[15] A. Peralta-Sevilla, E. Inga, R. Cumbal, and R. Hincapié, “Optimum Deployment of FiWi Networks using Wireless Sensors based on Universal Data Aggregation Points,” Commun. Comput. (COLCOM), 2015 IEEE Colomb. Conf., 2015.
[16] A. Alhamwi, W. Medjroubi, T. Vogt, and C. Agert, “GIS-based urban energy systems models and tools: Introducing a model for the optimisation of flexibilisation technologies in urban areas,” Appl. Energy, vol. 191, pp. 1–9, 2017.
[17] P. Nageler *et al.*, “Novel validated method for GIS based automated dynamic urban building energy simulations,” Energy, vol. 139, pp. 142–154, 2017.

[18] B. Howard, L. Parshall, J. Thompson, S. Hammer, J. Dickinson, and V. Modi, “Spatial distribution of urban building energy consumption by end use,” Energy Build., vol. 45, pp. 141–151, 2012.
[19] H. Li, W. Mao, A. Zhang, and C. Li, “An improved distribution network reconfiguration method based on minimum spanning tree algorithm and heuristic rules,” Int. J. Electr. Power Energy Syst., vol. 82, pp. 466–473, 2016.
[20] C. Mateo *et al.*, “European representative electricity distribution networks,” Int. J. Electr. Power Energy Syst., vol. 99, no. January, pp. 273–280, 2018.
[21] E. Inga, J. Inga, E. Correa, and R. Hincapié, “Reconstrucción del patrón de consumo eléctrico a partir de Big Data mediante técnica de MapReduce (Reconstruction of the Electric Consumption Pattern from Big Data using MapReduce Technique),” Enfoque UTE, vol. 9, pp. 177–187, 2018.
[22] J. Inga-ortega, E. Inga-ortega, and C. Gómez, “Electrical Load Curve Reconstruction required for Demand Response using Compressed Sensing Techniques,” IEEE PES Innov. Smart Grid Technol. Conf. - Lat. Am. (ISGT Lat. Am., 2017.
[23] G. Davidescu, T. Stutzle, and V. Vyatkin, “Network planning in smart grids via a local search heuristic for spanning forest problems,” IEEE Int. Symp. Ind. Electron., pp. 1212–1218, 2017.
[24] N. Funabikiy, K. Uemura, T. Nakanishi, and W. Hassan, “A minimum-delay routing tree algorithm for access-point communications in Wireless Mesh Networks,” 2008, pp. 161–166.
[25] B. Karimi, V. Namboodiri, and M. Jadliwala, “On the scalable collection of metering data in smart grids through message concatenation,” 2013, pp. 318–323.
[26] J. M. Smith and B. Toppur, “Euclidean Steiner minimal trees, minimum energy configurations, and the embedding problem of weighted graphs in E3,” Discret. Appl. Math., vol. 71, no. 1–3, pp. 187–215, Dec. 1996.
[27] B. Krishnamachari, D. Estrin, and S. Wicker, “The Impact of Data Aggregation in Wireless Sensor Networks,” Distrib. Comput. Syst. Work. 2002. Proceedings. 22nd Int. Conf., pp. 0–3, 2002.
[28] R. Xie and X. Jia, “Transmission-Efficient Clustering Method for Wireless Sensor Networks Using Compressive Sensing,” IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 3, pp. 806–815, Mar. 2014.
[29] I. F. Senturk, K. Akkaya, and F. Senel, “An effective and scalable connectivity restoration heuristic for Mobile Sensor/Actor Networks,” in 2012 IEEE Global Communications Conference (GLOBECOM), 2012, pp. 518–523.
[30] O. Kahveci, T. J. Overbye, N. H. Putnam, and A. Soylemezoglu, “Optimization framework for topology design challenges in tactical smart microgrid planning,” 2016 IEEE Power Energy Conf. Illinois, PECI 2016, pp. 1–7, 2016.
[31] H. L. Nguyen and U. T. Nguyen, “Bandwidth efficient multicast routing in multi-channel multi-radio wireless mesh networks,” in Ultra Modern Telecom-

munications & Workshops, 2009. ICUMT '09. International Conference on, 2009, pp. 1–8.
[32] S. Abeysinghe, J. Wu, M. Sooriyabandara, M. Abeysekera, T. Xu, and C. Wang, "Topological properties of medium voltage electricity distribution networks," Appl. Energy, vol. 210, pp. 1101–1112, 2018.
[33] C. Long, J. Wu, L. Thomas, and N. Jenkins, "Optimal operation of soft open points in medium voltage electrical distribution networks with distributed generation," Appl. Energy, vol. 184, pp. 427–437, 2016.
[34] K. G. Boroojeni, M. H. Amini, S. S. Iyengar, M. Rahmani, and P. M. Pardalos, "An economic dispatch algorithm for congestion management of smart power networks: An oblivious routing approach," Energy Syst., vol. 8, no. 3, pp. 643–667, 2017.
[35] M. H. Amini, B. Nabi, and M.-R. Haghifam, "Load Management Using Multi-Agent Systems in Smart Distribution Network," IEEE Power Energy Soc. Gen. Meet., pp. 1–5, 2013.

Sobre los autores



Esteban Mauricio Inga Ortega (Y'1976-M'08). Esteban Inga es investigador científico en la Universidad Politécnica Salesiana – Sede Quito. En el año 2001 se gradúa de Ingeniero Electrónico en la Universidad Politécnica Salesiana – Sede Cuenca. En el año 2008 recibió su Maestría en Educación y Desarrollo Social por la Universidad Tecnológica Equinoccial, a inicio del año 2017 recibe su grado de Magíster en Ingeniería por la Universidad Pontificia Bolivariana de Medellín – Colombia y en Noviembre de 2017 obtiene su grado de Ph.D de la Universidad Pontificia Bolivariana de Medellín – Colombia con mención de honor Magna Cum Laude otorgado cuando la tesis supera las expectativas de lo planteado en el proyecto, por concepto unánime de los jurados, y se demuestra producción científica. Es coordinador del grupo de investigaciones en redes eléctricas inteligentes GIREI de la Universidad Politécnica Salesiana, coordinador del programa de Maestría en Electricidad-Sede Quito, miembro activo de IEEE, investigador acreditado Senescyt y miembro del comité científico INEDITA-Senescyt



Miguel Campaña (Y'1989-M'05). Se graduó de Ingeniero Eléctrico en la Universidad Politécnica Salesiana-Ecuador, investigador del Grupo de Investigación GIREI. Su trabajo se basa en modelación matemática y técnicas de simulación de redes de distribución eléctrica. Sus intereses de investigación incluyen la asignación de recursos en redes de malla inalámbricas para AMI y el sistema de distribución de eléctrica soterrada y aérea.

Capítulo II

Métodos de localización de fallas en sistemas eléctricos de distribución con presencia de generación distribuida

Fault Location Methods in Electric Distribution Systems with Presence of Distributed Generation

Ramón Pérez* Carmen Vásquez§ Amelec Viloria†
*email: rperezp @ups.edu.ec
§email: cvasquez @unexpo.edu.ve
†email: avioloria7@cuc.edu.co

Resumen

La localización de fallas en los sistemas eléctricos, especialmente en los de distribución, es un problema que afecta no solo a los usuarios, sino también a las empresas prestadoras del servicio eléctrico; con mayor tiempo invertido en este período, se incrementa las pérdidas por la energía no facturada y las molestias en los usuarios, disminuyendo la calidad del servicio. Son distintas las causas por las cuales se incrementa este tiempo, entre éstas la confusión que pudiera existir en los sistemas de localización que actúan bajo la presencia de fuentes de generación distribuida en las redes de distribución. En este sentido, la presente investigación muestra una revisión de los métodos utilizados para la localización de fallas en sistemas eléctricos de distribución que tengan presencia de fuentes de generación distribuida. Las principales características de los métodos son estudiadas y sus más recientes aplicaciones para tratar este problema.

Palabras clave

Fallas, inteligencia artificial, generación distribuida, sistemas eléctricos de distribución.

Abstract

The location of faults in the electrical systems, especially in distribution, is a problem that affects not only the users, but also the companies providing the electric service; with more time invested in this period, the losses due to the unbilled energy and the inconvenience in the users are increased, decreasing the quality of the service. The causes for which this time is increased are different, among them the confusion that could exist in the location systems that act under the presence of sources of distributed generation in the distribution networks. In this sense, the present investigation shows a revision of the methods used for the location of faults in electrical distribution systems that have presence of sources of distributed generation. The main characteristics of the methods are studied and their most recent applications to deal with this problem.

Keywords

Faults, artificial intelligence, distributed generation, electric distribution systems.

1. Introducción

Un Sistema Eléctrico de Distribución (SED) es un conjunto de instalaciones y equipos encargados de entregar la energía eléctrica a los distintos usuarios. En general, existen tres (3) tipos de configuraciones para este tipo de sistema y que se suelen llamar radial, anillo y malla, cada una con sus características, ventajas y desventajas. La configuración radial es más económica y sencilla de operar, es la más común para alimentar cargas de tipo residencial y no críticas [1]. Adicionalmente, según su construcción, estos sistemas se clasifican en aéreos y subterráneos, siendo los aéreos los más comunes [2].

Los SED no son inmunes a las interrupciones del servicio eléctrico ocasionadas por fallas. El tiempo de una interrupción depende de la detección de la falla por el dispositivo de protección, su apertura y despeje, de la localización de la falla y de la reparación necesaria para restablecer el servicio. La localización de la falla es el punto crítico del proceso y el principal inconveniente está relacionado con aspectos de infraestructura, es decir, la mayoría de los circuitos en los SED no disponen de localizadores de fallas, la configuración de los circuitos no es homogénea, ya que generalmente contienen diferentes secciones de conductor, longitudes y presentan múltiples derivaciones de cargas a lo largo de cada alimentador [3].

Actualmente, los SED están presentando penetración de Generación Distribuida (GD), la cual es universalmente aceptada como una solución efectiva y económica para responder a los incrementos en las demandas de energía del sistema, debido a que representa una mejor opción para corregir los problemas de pérdidas de energía, perfiles de tensión, problemas de cargabilidad de líneas, seguridad y confiabilidad, por mencionar algunos [4]. La integración de la GD a los SED convencionales modifica las amplitudes de las señales de falla (tensión y corriente) lo cual afecta significativamente la exactitud de los algoritmos de localización de fallas [5], por esta razón, este problema resulta de importancia tratarlo cuando existe presencia de GD en las redes [6].

Los métodos algorítmicos han sido utilizados para tratar este problema con la presencia de GD. Estos métodos dependen del modelo de la red y el error en la localización de la falla aumenta significativamente cuando se incrementa la potencia inyectada por las fuentes de GD a los SED [7]. Los autores de [8] utilizan el método basado en impedancia para localizar fallas en SED con presencia de GD. La técnica es validada en el circuito IEEE de 34 nodos considerando los diferentes tipos de falla que se puedan presentar con resistencias de falla entre 0 y 40 Ω y penetración de GD entre 5 y 50%. Los resultados los muestran errores de estimación menores al 2%.

Los autores de [9] concluyen que los métodos basados en la impedancia están influenciados por la resistencia de falla y por la distancia entre la falla y el punto de medición. Mientras mayor sea la resistencia y la distancia de falla, mayor será el error en la estimación. En general, la exactitud de los métodos basados en la impedancia depende de los parámetros de la línea, sus características y del valor de carga. El error en la localización de estos métodos también se ve afectada por la complejidad de la red, como los sistemas no balanceados, múltiples laterales y resistencias de falla. Para fallas con múltiples estimaciones, el estado de los dispositivos de protección se utiliza comúnmente para identificar la ubicación real. Sin embargo, para un SED que no está equipado con el estado en línea de los dispositivos de protección, el problema de la múltiple estimación no podría resolverse [10]. Estos métodos proporcionan exactitud, pero con múltiple estimación en la localización de la falla [11].

En [12] localizaron fallas en un SED con presencia de GD utilizando Redes Neuronales Artificiales (ANN, por sus siglas en inglés) con el enfoque estándar de retropropagación. En este método, los datos de entrenamiento se basan en la corriente inyectada por cada fuente de fuente de GD durante la falla. Por lo tanto, la exactitud del método es altamente dependiente del número de fuentes de GD en el sistema. El principal inconveniente de las ANN es su alta dependencia de la cantidad y la calidad de los datos entrenados para producir un algoritmo bien entrenado. Una cantidad limitada de información o su inexactitud, afecta el desempeño del algoritmo para identificar correctamente la ubicación de la falla. Este problema ocurre en los SED con información limitada que resulta de un número insuficiente de dispositivos de monitoreo [10].

Las Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) han sido utilizadas para localizar fallas en SED con GD. Éstas se basan en patrones representados por medidas de tensión y corriente en la subestación y en las fuentes de GD. Esta técnica presenta mayor robustez cuando se incrementa el número de fuentes de GD en la red. En [13] presentan la aplicación de las SVM para diagnosticar fallas en SED con la presencia de GD. El enfoque propuesto se basa en las tres (3) tensiones y corrientes de fase que están disponibles en todas las fuentes, es decir, en la subestación y en los puntos de conexión de la GD. La metodología propuesta es ilustrada en un alimentador de distribución de la subestación de 132/11kV en la India con cargas en diferentes lugares y varias fuentes de GD. El esquema de localización de fallas propuesto es capaz de identificar con exactitud el tipo de falla, la ubicación de la sección del alimentador en falla y la impedancia de falla. El resultado de la simulación muestra el rendimiento satisfactorio en términos de clasificación y regresión. La exactitud de la clasificación para la línea en falla fue de 100, 99,95 y 92,06%, con tres (3) fuentes de GD, dos (2) fuente de GD y una (1) fuente de GD, respectivamente. Para cualquier cambio en la topología del sistema, las SVM tiene que ser reentrenadas antes de su aplicación. Además, con un mayor número de fuentes de GD, el enfoque se hace más robusto. Diferentes esquemas de localización de fallas en SED han sido estudiados considerando los métodos tradicionales basados en el modelado matemático del sistema, en ondas viajeras de altas

frecuencias, en inteligencia artificial y la integración de dos (2) o más de éstos, resultando en nuevas investigaciones que buscan dar solución a este problema. En este sentido, en esta investigación se presenta la revisión de los principales métodos para la localización de fallas en SED con presencia de GD. La descripción de los SED, la GD y las metodologías existentes para tratar este problema se presentan en las siguientes secciones, finalizando con las conclusiones.

2. Sistemas eléctricos de distribución

Los Sistemas Eléctricos de Potencia (SEP) y la prestación del servicio eléctrico incluyen las actividades de generación, transmisión, distribución, despacho y comercialización final de la energía. Entre estas actividades, la de distribución es conocida como el "gigante invisible" [2], por ser aproximadamente el 66% de la inversión total del sistema, lo que implica que se debe realizar un trabajo riguroso a la hora de planificar, diseñar, construir, operar y mantenerlo. Su función primordial es llevar la energía eléctrica generada en las centrales a los usuarios finales y entregarla con los requisitos de calidad exigidos, es decir, de manera, segura, confiable y de adecuada prestación [14].

Los SED están constituidos por las subestaciones, conductores, transformadores y demás elementos de un sistema eléctrico que sirve para llevar la energía eléctrica desde la subestación distribuidora hasta los usuarios finales. Esta subestación incluye transformadores de potencia, barras, condensadores, seccionadores y reconectadores, entre otros elementos. Los alimentadores primarios de distribución, a la salida de las subestaciones, son trifásicos y operan a tensiones que van desde 1.000 V hasta 34.5 kV y se encargan de distribuir la energía a los centros de consumo. A partir de estos centros se tienen sub-alimentadores ramales trifásicos y/o monofásicos. Los transformadores de distribución pueden ser de diferentes tipos y configuraciones y sus potencias nominales van desde 10 a 500 kVA y suelen instalarse en los alimentadores primarios y sub-alimentadores para reducir el nivel de tensión de distribución en su circuito secundario al requerido por los usuarios [1]. Estos sistemas se clasifican de manera más general, en función de la disposición de sus

conductores y equipos, en aéreos y subterráneos. En la Figura 1 se muestra un SED típico.

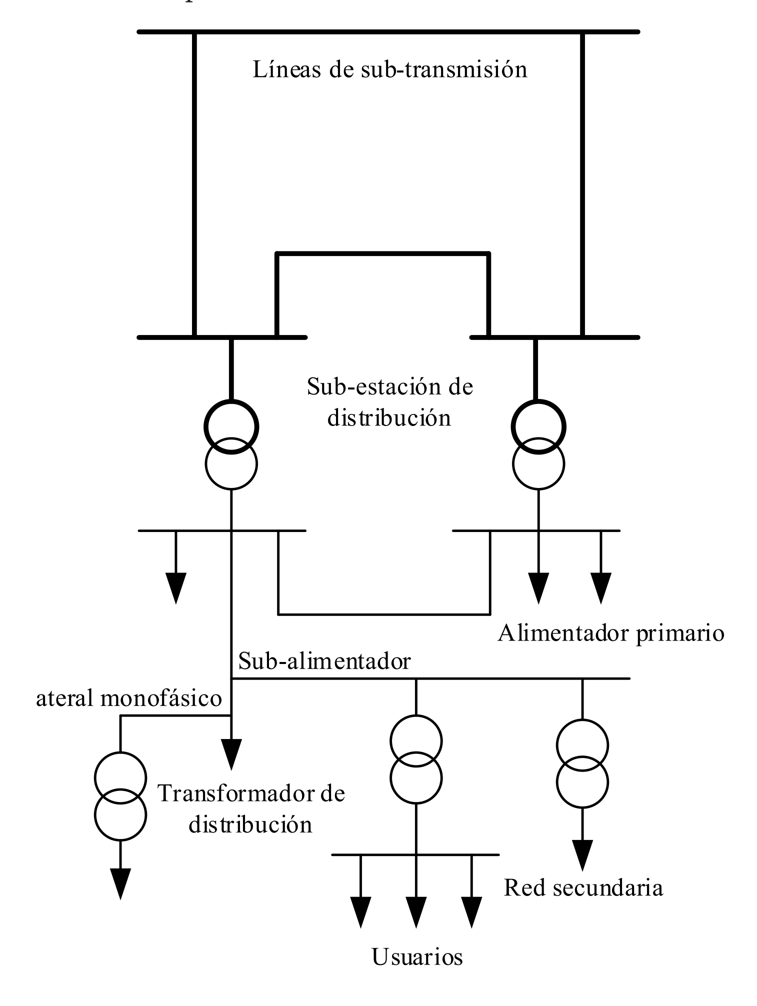


Figura 1. SED típico

Fuente: Adaptación de [15]

Según [2], las actividades de los SED están en constante evolución, por mencionar algunos de estos nos encontramos:

- Expansión de las cargas.
- Normalización de los materiales, estructuras y montajes.
- Técnicas de diseño, análisis y operación.
- Métodos de trabajos específicos y normalizados.
- Programas de prevención de accidentes y de mantenimientos.
- Surgimiento de nuevas empresas de fabricación de equipos.
- Grandes volúmenes de información, datos y planos.

Los SED son de grandes longitudes y número de equipos, por lo que los estudios para garantizar la calidad de la energía eléctrica prestada a sus usuarios son complejos de realizar, porque se involucran un sinnúmero de variables.

1) Interrupciones y las causas de fallas

Las interrupciones del servicio eléctrico tienen una multitud de origen o causas, por lo que es necesario realizar una clasificación y análisis de éstas con el fin de minimizarlas. Una primera clasificación que se puede hacer es en base a su naturaleza, según ésta las interrupciones pueden ser programadas o forzadas [11].

Las interrupciones programadas son previstas, ya que con suficiente antelación se les informa a los usuarios afectados, para que puedan tomar las distintas previsiones que minimicen el impacto causado por la falta del suministro de energía eléctrica. Este tipo de interrupciones suelen ocurrir por labores de mantenimiento, ampliaciones o maniobras de la red.

Las interrupciones forzadas son consideradas más perjudiciales, por su carácter imprevisto, debido a que los usuarios no tienen oportunidad de tomar medidas que el caso requiera. Según [16], estas últimas se clasifican de manera general según su origen el cual puede ser:

- Origen desconocido: Incluye todas las interrupciones cuyas causas se desconocen, suelen ser transitorias y no dejan algún rastro.
- Origen climatológico: Incluye todas las interrupciones que tienen su origen en fenómenos climatológicos, tales como lluvias o descargas atmosféricas.

- Origen externo: Incluye las interrupciones debidas a agentes externos al sistema eléctrico como son los animales (aves, roedores), árboles, ramas, vehículos, personas, por nombrar algunos.
- Origen interno: Incluye todas las interrupciones causadas por agentes o elementos pertenecientes al sistema eléctrico, entre los cuales destacan: falla en algún componente de la red, falla en los sistemas de protección, falla humana en el manejo del sistema, por nombrar algunos.

Según [17], otra forma de clasificar las interrupciones es de acuerdo con su duración, entre las cuales destacan:

- Momentáneas: Una interrupción momentánea se presenta cuando el valor eficaz de la tensión de suministro disminuye a un valor por debajo de 0,1 por unidad (pu) y durante un tiempo inferior a los tres ciclos (50 ms), pero superior o igual al medio ciclo (8,33 ms). No requiere la intervención humana para reparar y restablecer el normal funcionamiento del sistema
- Temporales: Una interrupción temporal se presenta cuando el valor eficaz de la tensión de suministro disminuye a valores por debajo de 0,1 pu, durante un tiempo inferior a un (1) min, pero superior o igual a los tres ciclos (50 ms). Tampoco requiere la intervención humana para reparar y restablecer el normal funcionamiento del sistema
- Permanentes o Sostenidas: Una interrupción permanente del servicio se origina cuando el valor eficaz de la tensión en el punto de suministro disminuye por debajo de 0.1 pu durante un tiempo superior a un (1) min. Este tipo de interrupción requiere la intervención humana para localizar, reparar y restablecer el normal funcionamiento del sistema y son las que mayor afectan la continuidad del servicio eléctrico.

Existe una clasificación más general para las interrupciones, que según su tiempo de duración pueden ser de corta o larga duración, en función de que la misma sea menor o mayor a un (1) minuto respectivamente. La tabla 1 muestra los requisitos establecidos por países europeos para establecer los límites de calidad. Adicionalmente, la tabla 2 muestra los datos de frecuencia y tiempo de duración promedio de las interrupciones de países de la misma región.

Tabla 1. Requisitos de las interrupciones en diferentes países europeos

País	Duración por evento (hr/evento)	Indisponibilidad (hr/año)	Número de eventos de larga duración	Número de eventos de corta duración
Bélgica	4			
Finlandia	12		12-24 (BT)	
			8-20 (MT)	
Francia	6		2-6 (AT)	2-30 (MT)
Italia	24		2-6 (BT)	
			3-5 (MT)	
Lituania		60 (BT)	1 (AT)	
Polonia				
Portugal	20-24	6-20 (BT)		
		4-16 (MT)		
		4 (AT)		
España	18 (BT)	6-20 (BT)		
		4-16 (MT)		
		6 (AT)		
Gran Bretaña	18 (MT)			
República Checa	12			
Hungría	12-18			

Tabla 2. Promedios de interrupciones del servicio eléctrico en países europeos durante el período 1996-2002

País	Frecuencia de la interrupción (evento por año)	Tiempo de Indisponibilidad del servicio eléctrico (min/año)
Australia	0,59	35,23
Bélgica	0,90	42,25
Finlandia	4,06	186,67
Francia	1,21	53,33
Alemania	0,27	57,00
Gran Bretaña	0,77	70,09
Irlanda	1,34	235,67
Italia	3,83	202,85
Noruega	2,73	218,00
España	2,98	153,00
Suecia	2,07	99,00
Países Bajos	0,38	27,50

Hay diversas causas de fallas en los SED, según estudios llevados a cabo por el Instituto de Investigación de Potencia Eléctrica (EPRI, por sus siglas en inglés) aproximadamente el 40% de fallas ocurren durante períodos de condiciones climáticas adversas que incluyen la lluvia, la nieve y el hielo [18].

3. Generación distribuida

Según [19], GD significa generadores de pequeña capacidad de potencia conectados directamente a los SED; su propósito es proporcionar una fuente de energía eléctrica activa [20]. Ésta es considerada como una opción importante para solucionar la mayoría de problemas en los SED, tales como pérdidas en las líneas, perfiles de tensión, cargabilidad de las líneas, confiabilidad y también alcanzar el objetivo de utilizar recursos energéticos verdes y renovables [4].

La capacidad máxima de GD que puede conectarse a un SED depende de su potencia, la cual está correlacionada con su nivel de tensión. El diseño técnico de cada SED es único, por lo tanto, no se puede dar una definición general de la capacidad máxima que se puede conectar, sin embargo, de acuerdo con el rango de potencia de la GD se sugieren algunas clasificaciones, tal como se muestra en la Tabla 3.

Tabla 3. Clasificación de la GD de acuerdo con su rango de potencia

Clasificación de la GD	Rango de potencia
Micro	~ 1 W < 5 kW
Pequeña	5 kW < 5 MW
Mediana	5 MW < 50 MW
Grande	50 MW < 300 MW

Fuente: [20]

De acuerdo con el tipo de tecnología disponible, la GD varía ampliamente, muchas de las tecnologías utilizan recursos de energía renovable. Estos se definen como recursos que generalmente no están expuestos al agotamiento, como la luz solar, la fuerza del viento, la materia orgánica (biomasa) y la del océano, por mencionar algunos. Otro tipo de fuentes de GD de acuerdo con su tecnología,

son las provenientes de la producción combinada de calor y electricidad. Las turbinas de gas de ciclo combinado, motores de combustión interna, turbinas de combustión y pilas de combustible son parte de este ciclo combinado. En la Tabla 4 se muestra los valores típicos de potencia de GD de acuerdo con su tecnología.

Tabla 4. Valores típicos de potencia de GD de acuerdo con su tecnología

Tecnología	Valor típico de potencia
Turbina de gas de ciclo combinado	35 – 400 MW
Motores de combustión interna	5 kW – 10MW
Turbinas de combustión	1 – 250 MW
Micro turbinas	35 kW – 1 MW
Renovables	
Micro hidroeléctricas	25 kW – 1 MW
Pequeñas hidroeléctricas	1 – 100 MW
Turbina eólica	200 W – 3 MW
Arreglos fotovoltaicos	20 W – 100 kW
Solar térmica	1 – 10 MW
Biomasa basada en gasificación	100 kW – 200 MW
Pilas de combustible, fosfato	200 kW – 2 MW
Pilas de combustible, carbonato fundido	250 kW – 2 MW
Geotérmica	5 – 100 MW
Energía del océano	100 kW – 1 MW
Baterías de almacenamiento	500 kW – 5 MW

Los SED se diseñan generalmente en un esquema radial, que tradicionalmente se alimenta desde la subestación ubicada aguas arriba en el inicio del circuito. Por lo tanto, se tiene un flujo de potencia unidireccional y una misma corriente dirigida desde la subestación hacia el alimentador aguas abajo, lo que facilita el cálculo de corriente de falla. Por otro lado, el nivel de tensión de las barras disminuye en la medida en que se encuentren más alejadas de la subestación, por esta razón, tienen un bajo perfil de tensión que debe ser corregido por bancos de condensadores, compensadores síncronos, Sistemas Flexibles de Transmisión en Corriente Alterna (FACTS, por sus siglas en inglés) o fuentes de GD en estos puntos débiles. Aunque la adición de GD en los SED mejora el perfil de

tensión, cambia drásticamente las características del sistema, particularmente en cuanto a la condición de falla y la respuesta de los dispositivos de protección, influyendo en los niveles de corriente y en la contribución de las fuentes [4].

4. Revisión sobre los métodos de localización de fallas

Los aspectos mayormente considerados por los métodos de localización de fallas en los SED están directamente relacionados con sus características, tales como heterogeneidad de los alimentadores (debido a los diversos tamaños y longitudes de los conductores), desbalances (debido a la presencia de cargas monofásicas y/o trifásicas) y la presencia de laterales a lo largo del alimentador principal, por mencionar algunos. Estos aspectos pueden introducir errores en la estimación de la localización de la falla si se usan modelos simplificados [11].

Los métodos de localización de fallas en SED pueden dividirse en tres (3) categorías, según [21]:

1. Métodos basados en la inspección visual.
2. Métodos convencionales.
3. Métodos basados en Inteligencia Artificial.

1) Métodos basados en la inspección visual

Inicialmente la localización de fallas en SED se realizaba de manera visual. Los dispositivos de protección en las subestaciones o el reclamo realizado por los usuarios del servicio a causa de una suspensión alertaban a cuadrillas de personal especializado a localizar las fallas de manera visual, para poder repararla y restablecer nuevamente el servicio. El tiempo total de indisponibilidad del servicio era función de la alerta de la falla, su localización y del tiempo utilizado para la reparación y restablecimiento, el cual podría variar de unos cuantos minutos a horas y días, dependiendo de diversos factores, entre estos, tipo de configuración (aérea o subterránea), de circuito (urbano o rural), de causa de la falla (agentes externos o internos al sistema), entre otros.

Los requerimientos de la calidad en la prestación del servicio hicieron que fuera cada vez más exigente reducir tiempos de indisponibilidad del servicio y para poder llevar un control de los mismos comenzaron a aparecer los indicadores, según las regulaciones establecidas [22], [23]. La norma IEEE Std 1366 [24] incluye una variedad de indicadores en función del número de clientes o usuarios afectados, la potencia interrumpida o la cantidad de energía dejada de suministrar, relativo a cuantificar los efectos relacionados a la suspensión del suministro eléctrico. Las iniciativas para reducir los tiempos de indisponibilidad del sistema de manera diferente a la inspección visual comenzaron a aparecer, como se presentan a continuación.

2) Métodos convencionales

Estos incluyen a los métodos basados en la impedancia y en ondas viajeras [21]

2.1.Métodos basados en la impedancia

Según [25], los métodos basados en la impedancia usan las mediciones de tensión y corriente de la frecuencia fundamental disponibles en la subestación e información como topología del SED, datos de líneas y cargas para encontrar la localización de la falla. A diferencia de las líneas de transmisión, los alimentadores de distribución suelen tener cargas intermedias significativas, mientras que las mediciones de tensión y corriente sólo están disponibles en la subestación y/o en la fuente de GD. Por lo tanto, estos métodos inician el proceso de localización de fallas desde la primera sección de línea y resuelven iterativamente las ecuaciones que describen la condición de estado de falla para todas las secciones, una por una, para hacer una estimación de la distancia. En la Figura 2 se ilustra este método.

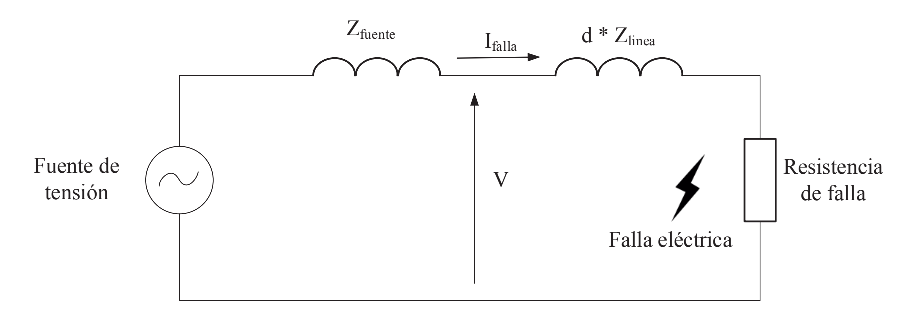


Figura 2. Método de localización de falla basado en la impedancia

La formulación de la distancia de ocurrencia de la falla se puede expresar por medio de la ecuación (1):

$$d = \frac{V}{I * Z_l} \tag{1}$$

donde:

V: Tensión durante la falla en V.
I: Corriente durante la falla en A.
Z_l: Impedancia de la línea en Ohm por unidad de longitud
d: Distancia de la falla (m)

Anteriormente, los métodos basados en la impedancia no consideraban la presencia de la GD [26]–[28]. Se han desarrollado algunas técnicas basadas en la impedancia para considerar la presencia de GD. Inicialmente estos métodos se basaban en la modificación o re-coordinación de los dispositivos de protección existentes en el SED y el objetivo era la identificación de la zona que presentaba la falla, pero sin una localización exacta de la misma [29]. Para determinar la localización exacta de la falla se han considerado diferentes parámetros, entre éstos variaciones en la tensión medida en la subestación y en las fuentes de GD antes y durante la falla [30], [31], uso de las componentes simétricas [32]–[35], entre otros. En la Tabla 5 se muestran los resultados de aplicación de los

métodos basados en la impedancia para la localización de fallas en SED con GD.

Sin embargo, cuando un SED tiene diversos laterales con presencia de fuentes de GD, la exactitud de este método es cuestionable [36]. El hecho que un alimentador tenga diversos laterales aumenta la complejidad de localizar la falla. Esto se debe principalmente a que la estimación de la localización de la falla produce más de una ubicación [37]. De igual manera, el modelo de línea utilizado en el SED afecta la exactitud de estos métodos, en el caso de líneas de distribución cortas, las capacidades de derivación son despreciables y el modelo de línea corta es suficiente. Sin embargo, para líneas aéreas o subterráneas largas, el efecto de la admitancia de derivación es mayor y puede resultar en errores de estimación si no se modela correctamente. En general, un correcto modelado de las líneas del SED mejorará la exactitud de los resultados. Por otro lado, modelos detallados aumentarían la complejidad y la carga computacional [25]. En la Tabla 5 se muestran los resultados de aplicación de los métodos basados en la impedancia para la localización de fallas en SED con GD.

2.2.Métodos basados en ondas viajeras

Este método se basa en el principio de transmisión y reflexión de las ondas que se producen entre el terminal de línea y el punto de falla. En la Figura 3 se ejemplifica el método de las ondas viajeras. Cuando se produce una falla, se generan ondas de alta frecuencia de corrientes y tensiones que se propagan hacia ambos extremos. En los puntos de discontinuidad, como circuito abierto, cortocircuito o terminales de línea, una parte de la onda incidente se refleja hacia atrás mientras otra parte viaja hacia más allá de la discontinuidad. Las ondas se reflejan hacia adelante y atrás entre el punto de falla y los dos (2) terminales hasta que se alcanza el estado estacionario post-falla [25].

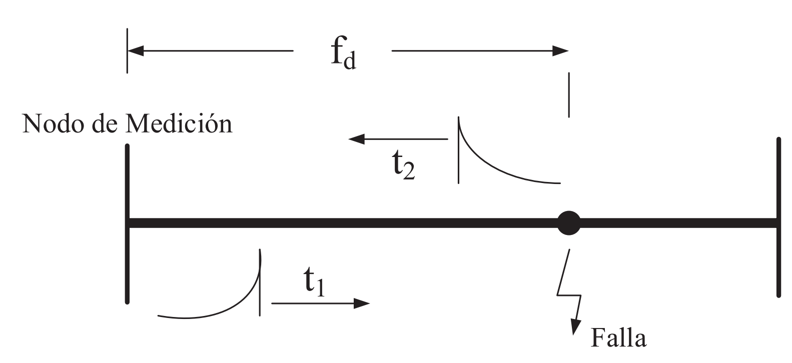


Figura 3. Método de las ondas viajeras

Fuente: [21]

En la Figura 3, t_1 es el tiempo que tarda la onda viajera en llegar al punto de falla desde el nodo de medición, t_2 es el tiempo que tarda la onda reflejada desde el punto de falla al nodo de medición y f_d es la distancia de falla medida desde el punto de medición. La distancia de falla se determina usando (2)

$$f_d = \frac{v * (t_2 - t_1)}{2} \qquad (2)$$

Para cada localización de falla, los transitorios registrados contienen algunas frecuencias características dominantes en función de la longitud de las trayectorias de desplazamiento y de las velocidades de propagación que proporcionan información valiosa sobre su ubicación. Sin embargo, la interpretación de los transitorios, especialmente para SED ramificados, son computacionalmente difíciles [25]. Por lo tanto, los métodos basados en ondas viajeras se soportan principalmente en técnicas de procesamiento de señal tales como transformaciones de onda para simplificar el análisis, entre éstas se encuentra la transformada de Wavelet, la cual es una transformación lineal similar a la transformada de Fourier, que permite una representación tiempo-frecuencia de la señal. Esta propiedad es particularmente útil para la localización temporal de transitorios de falla que tienen componentes de alta frecuencia y de corta duración superpuestos a formas de onda continuas de frecuencia de potencia.

El método de ondas viajeras requiere de dispositivos de adquisición de datos de alta velocidad, sensores, detectores de transitorios de falla y GPS para capturar la forma de onda transitoria para localizar la falla. La ventaja de este método es que no se ve afectada por la variación de la carga ni por altos valores de resistencia de conexión a tierra [21].

El método de ondas viajeras ha sido aplicado para la localización de fallas en SED con GD, basado en la transformada de Wavelet [38], [39] y en mediciones múltiples en el sistema [40]. En la Tabla 6 se muestran los resultados de aplicación de los métodos basados en ondas viajeras para la localización de fallas en SED con GD.

3) Métodos basados en Inteligencia Artificial

Debido a la complejidad de los SED y los diversos factores de incertidumbre, como la característica de los conductores (longitud y calibre) y el valor de la resistencia de falla, los cuales son difíciles de abordar con los métodos basados en la impedancia y las ondas viajeras, han aparecido métodos basados en el conocimiento para la localización de fallas en estos sistemas. En general, estos métodos requieren de patrones que caractericen la falla tales como las mediciones de tensión y corriente.

La Inteligencia Artificial (IA) es un subcampo de la informática que investiga cómo el pensamiento y la acción de los seres humanos pueden ser imitadas por las máquinas [41]. Tanto los cálculos numéricos, no numéricos y simbólicos están incluidas en el ámbito de la IA. La imitación de inteligencia incluye no sólo la capacidad de tomar decisiones racionales, sino también para hacer frente a información con datos faltantes, adaptarse a las situaciones existentes y mejorarse a sí misma en el horizonte a largo plazo sobre la base de la experiencia acumulada. Estos métodos basados en IA para localizar fallas en SED se mencionan a continuación.

3.1.Redes neuronales

Una Red Neuronal Artificial (ANN, por sus siglas en inglés) es un conjunto de neuronas primarias que están conectados juntas en

diferentes arquitecturas organizadas en capas que están inspirada en la estructura del sistema nervioso de los seres humanos [41]. Las ANN son diferentes de los sistemas expertos, ya que no necesitan una base de conocimiento para trabajar. En su lugar, tienen que ser entrenadas con numerosos casos reales. Las ANN es uno de los métodos conocidos y confiables en la solución de problemas de ingeniería relacionados con la clasificación y optimización y por la capacidad que tienen en el reconocimiento de patrones complejos, su uso se ha extendido a la localización de fallas en SED. En la Figura 4 se presenta la estructura de una ANN feedfoward conocida como perceptron.

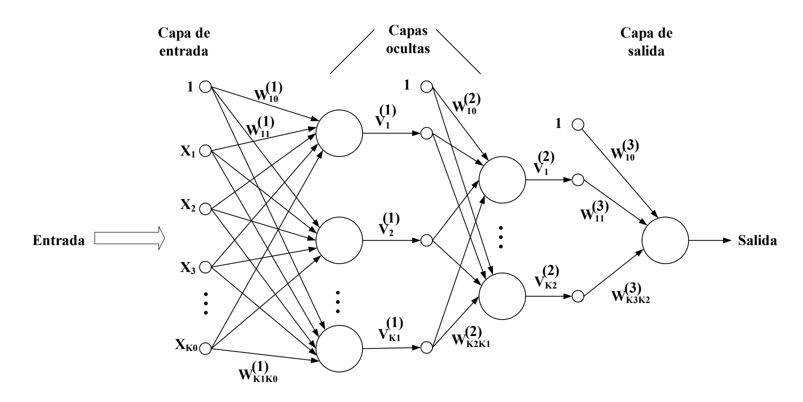


Figura 4. Representación gráfica de las ANN

Fuente: [41]

De la Figura 4 se distingue una capa de entrada la cual está compuesta por neuronas que reciben datos o señales procedentes del entorno. La capa de salida que está compuesta de neuronas que proporcionan la respuesta de la red neuronal. Una capa oculta que no tiene conexión directa con el entorno y proporciona grados de libertad a la red neuronal lo que le permite representar de manera más aproximada determinadas características del entorno que trata de modelar. Los datos primero ingresan a la capa de entrada luego pasan por la capa oculta y finalmente el resultado se obtiene en la capa de salida. Para la localización de fallas en sistemas eléctricos de distribución, los datos de entrada de la ANN pueden ser mediciones

realizadas en el circuito, tales como tensión, corriente, el estado de los dispositivos de protección, entre otros. La salida de la ANN corresponde a la ubicación de la falla. En la Tabla 7 se presentan los resultados de la aplicación de las ANN en la localización de fallas en SED con GD.

La localización de fallas basada en ANN sufre de inexactitudes debido a las variaciones continuas de carga e impedancia de la fuente [41]. De igual manera, es altamente dependiente de la cantidad y calidad de los datos entrenados para producir un adecuado algoritmo, por lo tanto, una cantidad limitada de información afecta la eficiencia del método. Este problema ocurre en los SED con información limitada resultante de un número insuficiente de dispositivos de monitoreo. Otra desventaja de las ANN es que el proceso de entrenamiento tiene convergencia lenta. Además, los parámetros como las capas ocultas, las neuronas y la tasa de aprendizaje se identifican utilizando casos de prueba y error. De igual manera, las ANN deben ser reentrenadas cada vez que el SED experimente cambios [21].

3.2.Máquinas de soporte vectorial

Las bases de esta herramienta fueron desarrolladas por Vladimir Vapnik [42] y son populares debido a muchas características atractivas y prometedoras que presenta. La formulación incorpora la Minimización del Riesgo Estructural (SRM, por sus siglas en inglés) que se ha demostrado ser superior al principio de Minimización del Riesgo Empírico (ERM, por sus siglas en inglés), empleado por las ANN convencionales [43]. La SRM minimiza un límite superior en el riesgo esperado, en contraposición a ERM que minimiza el error en los datos de entrenamiento. Es esta diferencia la que equipa las SVM con una mayor capacidad de generalizar, que es el objetivo en el aprendizaje estadístico. Las SVM se han desarrollado para resolver el problema de clasificación y regresión [44].

El concepto de las SVM se muestra en la Figura 5 para la clasificación entre dos (2) clases (clase 1 y clase 2). Los puntos se marcan usando el conjunto de entrenamiento de los datos de tensión y/o corriente tomados en el nodo de medición para la clasificación de fallas. Los vectores de soporte son los elementos de los conjuntos

de entrenamiento, que identifican el Hiperplano de Separación Óptima (OSH, por sus siglas en inglés). Los círculos representan la clase 1 y los cuadrados la clase 2. La entrada para las clases 1 y 2 son del nodo de medición (tensión y/o corriente) y la salida es la ubicación de la falla. El OSH separa todas las muestras y también maximiza el margen entre los dos (2) planos paralelos que contienen el conjunto de clases. El margen se define como la suma de la distancia mínima entre el conjunto de datos de entrenamiento y el OSH. En la figura 5 se muestra la representación gráfica de las SVM. De (3) a (6) se muestran las ecuaciones de los planos paralelos al OSH, el margen y la distancia al origen.

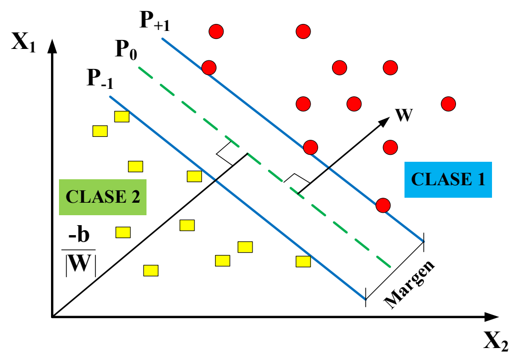


Figura 5. Representación gráfica de las SVM

$$P_{+1}: (\vec{w}.\vec{x_i}) + b = +1 \quad (3)$$

$$P_{-1}: (\vec{w}.\vec{x_i}) + b = -1 \quad (4)$$

$$\frac{\vec{w}}{\|\vec{w}\|}.(\overrightarrow{x_{+1}} - \overrightarrow{x_{-1}}) = \frac{2}{\|\vec{w}\|} \quad (5)$$

$$\frac{(\vec{w}.\vec{x_i}) + b}{\|\vec{w}\|} = \frac{b}{\|\vec{w}\|} \quad (6)$$

Para encontrar el OSH se debe maximizar el margen (5), considerando la restricción de que no existan datos dentro del margen, lo q

$$\min_{w} \frac{1}{2}(\vec{w}.\vec{w}) \quad (7)$$

$$\text{sujeto a } y_i(\vec{w}.\vec{x_i} + b) \geq 1, \forall i \quad (8)$$

La ecuación (7) se llama función objetivo y, junto con (8), representa el problema de optimización cuadrático con restricciones, lo que puede ser resuelto utilizando el método de los multiplicadores de Lagrange, denotados por la ecuación (9)

$$L(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\|\vec{w}\|^2 - \sum_{i=1}^{n} \alpha_i[y_i(\vec{w}.\vec{x_i} + b) - 1] \quad (9)$$

Para el caso de que existan intersecciones en los datos de entrenamiento (datos no separables), tal como se muestra en la Figura 6, los puntos que se encuentran en la intersección no pueden ser clasificados correctamente y sus correspondientes tenderán al infinito.

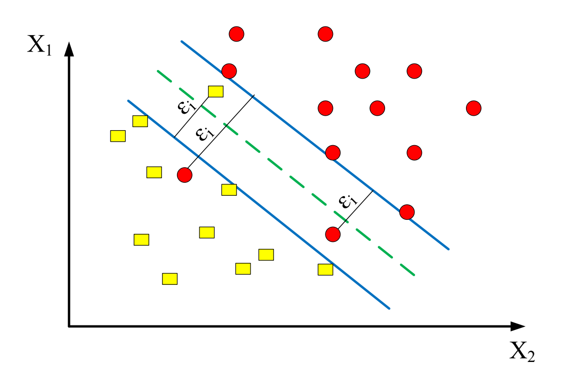


Figura 6. Hiperplano clasificador óptimo para caso no separable

En [42] introducen las variables slack o de relajación para permitir datos que violen las restricciones. De esta manera para obtener el OSH corresponde a minimizar la expresión (10)

$$\min_{w} \left[\frac{1}{2}(\vec{w}.\vec{w}) + C\sum_{i=1}^{n} \xi_i\right] \quad (10)$$

$$\text{Sujeto a } y_i(\vec{w}.\vec{x_i} + b) \geq 1 - \xi_i, \forall i$$

El parámetro C se elige por conveniencia por el usuario, de tal manera que, un valor grande sea una alta penalización a los errores. El valor óptimo de C se obtiene por medio de una validación cruzada [45].

Cuando los datos no son separables, se mapea el espacio de entrada a un espacio de representación de una dimensión mayor a través de una función no lineal elegida por conveniencia [46], tal como se muestra en la figura 7.

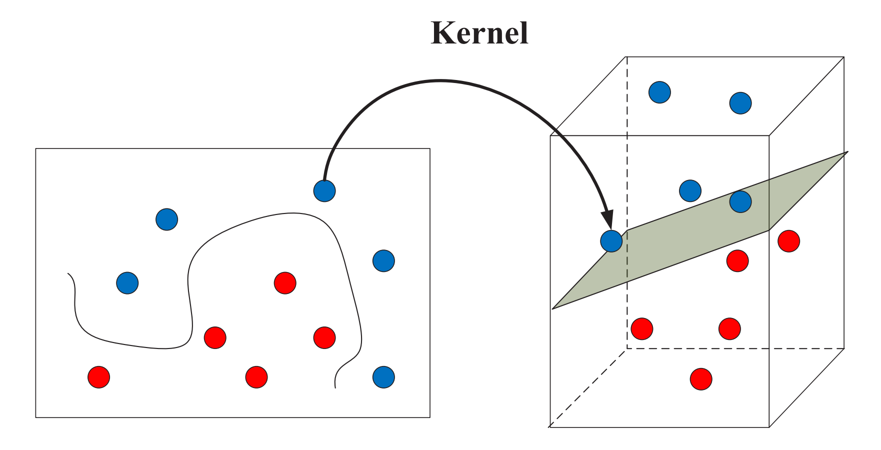


Figura 7. Transformación del espacio de entrada al espacio característico

Por medio de una función (Φ) se trazan los datos de entrada a algún espacio de mayor dimensión donde se puede manejar una separación lineal. Este espacio es llamado espacio característico (F) y es denotado por la relación (11).

$$\Phi: R^N \rightarrow F \tag{11}$$

La función (Φ) se define como el producto punto de los vectores en el espacio característico tal como se muestra en la ecuación (12).

$$k(\vec{x}_i, \vec{x_j}) = \Phi(\vec{x_i}).\Phi(\vec{x_j}) \tag{12}$$

La ecuación (12) se conoce también como la función Kernel (k) y entre las más usadas se encuentran la lineal, la RBF, la polinomial y la sigmoide. Estas funciones se muestran en la Tabla 8.

Tabla 8. Funciones Kernel más usadas

Kernel	Función
Lineal	$K(X_i,X_j) = X_i^T X_j$
RBF	$K(X_i,X_j) = exp(-\gamma\|\|X_i-X_j\|\|^2),\gamma>0$
Polinomial	$K(X_i,X_j) = (\gamma X_i^T X_j+r)^d,\gamma>0$
Sigmoide	$K(X_i,X_j)=tanh(\gamma X_i^T X_j+r)$

γ,r y d son los parámetros de la función kernel.
Fuente: [45]

Una SVM es un algoritmo de aprendizaje utilizado típicamente para resolver problemas de clasificación binaria y luego se extendieron para tratar problemas con múltiples clases. Estas múltiples clases son las que se presentan en la localización de fallas en SED y se resuelve por medio del esquema propuesto por [47]. En la Tabla 9 se presentan los resultados de la aplicación de las SVM en la localización de fallas en SED con GD.

4. Métodos integrados

Se dan tres (3) tipos de métodos para localizar las fallas en SED conocidos como la inspección visual, el basado en la impedancia y las ondas viajeras y los métodos inteligentes. Sin embargo, éstos todavía presentan algunos problemas, por ejemplo, el método basado en ondas viajeras puede enfrentarse a problemas tales como alta frecuencia de muestreo y estructura compleja, los métodos inteligentes pueden ser problemáticos debido a la necesidad de un banco de datos grande y detallado [48]. Finalmente, los métodos basados en la impedancia presentan problemas de múltiple estimación como se menciona en apartados anteriores.

El siguiente método consiste en la combinación de más de un método de localización de fallas y permite aprovechar la fortaleza de cada uno de éstos y así se podría obtener una mejor exactitud, este se conoce como método integrado [10]. En la Tabla 10 se muestra los métodos integrados para la localización de fallas en SED. Cabe destacar que tales combinaciones, además de sus ventajas, también traen desventajas. Si bien, la ventaja común de los mencionados

métodos integrados es su capacidad para superar el problema de estimación múltiple, tienen requerimientos adicionales en comparación con los algoritmos basados en impedancia tales como datos de entrenamiento, mediciones con mayor frecuencia de muestreo, medidas dispersas de tensión, por mencionar algunos. Para finalizar, en la Tabla XI se presenta una comparación de los diferentes métodos propuestos de localización de fallas en SED con GD.

5. Conclusiones

La localización de fallas en SED con GD es una problemática que ha sido abordada con diferentes técnicas, las cuales tienen sus ventajas y desventajas. Estas técnicas abarcan la inspección visual, los métodos convencionales, la inteligencia artificial y los métodos integrados. En esta investigación se presenta una revisión de los principales métodos utilizados para tratar el problema. Cada uno de ellos presenta sus ventajas y desventajas con errores de estimación aceptables. Se evidencia una tendencia a integrar los métodos existentes para mejorar los resultados. La presencia de GD en los SED extiende las investigaciones del tema debido a los efectos directos que tiene sobre las metodologías convencionales que daban solución al problema de localización de fallas en SED.

Tabla 5. Resultados de aplicación de los métodos basados en la impedancia para la localización de fallas en SED con GD

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
Impedancia aparente de secuencia positiva	1	Simulación de 67 fallas, consideración de cargas balanceadas y desbalanceadas. La GD de 3.2 MVA y 440 V interconectada a través de un transformador. Solo considera modelo de línea corta.	13.8 kV, 14 barras, 14.5 MVA. Alimentador radial sin presencia de laterales.	No indica.	Tensión y corriente de cada fase medida en la subestación.	Máximo error de 0.2% de la longitud total del alimentador para cargas balanceadas Máximo error de 0.2% de la longitud total del alimentador para cargas desbalanceadas	[32]such as the distribu≤ed generation (DG
Método general para localización de fallas en SED	6	Uso de equivalente de Thevenin para secuencia positiva, negativa y cero de las fuentes de GD. Primero se identifica la sección en falla y luego la ubicación de la falla. La GD representa el 49% del total de la carga del sistema. Se simularon fallas al 5, 50 y 95% de cada línea.	12.47 kV, 60 barras, 2.5 MVA de carga total en el sistema	1, 5, 10, 20, 25 y 50.	Tensión y corriente medida en las fuentes de GD y en la subestación.	Error de 3.28% para 4 fuentes de GD. 1.64% para 5 fuentes de GD. 0% para 6 fuentes de GD	[30]

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
Método de localización de falla de fases para redes de distribución con generación distribuida	1	Se considera la división del circuito de distribución en dos secciones: la primera para fallas ubicadas antes de la fuente de GD y la segunda para fallas ubicadas después de ña fuente de GD. Se presenta un esquema de localización de fallas para ubicaciones antes de la fuente de GD. Para ubicaciones posteriores el enfoque tratado son análisis de circuitos. Se considera el circuito de distribución balanceado y desbalanceado.	Nueve (9) barras. Alimentador radial sin presencia de laterales.	De 0.01 a 50.	Tensión y corriente medida en la subestación.	Error máximo de 0.18% para ubicación de fallas antes de la fuente de GD. Error máximo de 0.23% para ubicación de fallas después de la fuente de GD. Ambos para circuito balanceado	[49]
						Error máximo de 0.20% para ubicación de fallas antes de la fuente de GD. Error máximo de 0.30% para ubicación de fallas después de la fuente de GD. Ambos para circuito desbalanceado	
Modificación de los métodos basados en la impedancia	1	Solo se consideran fallas monofásicas y trifásicas. La fuente de GD suministra el 26% de la potencia activa total del sistema. Se simulan 500 casos de falla	11 kV, 27 barras,	0, 5, 10, 25 y 50.	Tensión y corriente medido en la subestación y en la fuente de GD.	El error varía entre -0.67% y 0.53% para fallas monofásicas y entre -0.2% y 0.12% para fallas trifásicas.	[50]

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
Modificación de los métodos basados en la impedancia	2	Se identifica la localización de la falla en dos (2) pasos: Primero se identifica todas las posibles ubicaciones de la falla y luego se determina la localización exacta de la falla basándose en la coincidencia de la tensión medida en la subestación y en cada fuente de GD con los valores calculados. El circuito de prueba es desbalanceado y considera el modelo de línea Pí.	24.9 kV, 34 barras	0, 10 y 25.	Tensión y corriente medido en la subestación y en la fuente de GD.	Máximo error de 0.1% para resistencia de falla de 25 Ohmios y distancia de falla de 50 km.	[29]

Tabla 6. Resultados de aplicación de los métodos basados en ondas viajeras para la localización de fallas en SED con GD

Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
1	La técnica propuesta extrae las componentes de alta frecuencia iniciadas por fallas de las señales de tensión, que se registran sólo en la subestación, utilizando la transformada wavelet	Circuito considera solo el alimentador principal dividido en 4 secciones. Solo considera modelo de línea corta. No considera ramales	No indica	Tensión medida en la subestación	No indica	[38]
2	Este método adquiere la información recogida por puntos de medición en las barras de la red de distribución y en los terminales del alimentador.	IEEE de 34 nodos	No indica	Mediciones simples y múltiples de tensión	Máximo error de 1.1545 km	[40]

Tabla 7. Resultados de aplicación de las ANN en la localización de fallas en SED con GD

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
MLPNN	1	Fallas de fase y fallas fase-tierra. Simulación de fallas en 67 ubicaciones diferentes dentro del circuito.	13 barras, 13.8 kV, 12 líneas y 11 cargas.	0, 10, 20,50 y 100.	Tensiones y corrientes trifásicas de la onda fundamental. Tercer armónico de las corrientes trifásicas. Componentes simétricas del segundo, tercero y quinto armónico de la corriente trifásica.	Menores al 5%	[51]
MLPNN	2	Fallas monofásicas y fallas trifásicas	27 barras, 20 kV, 13 km de longitud.	No indica.	Corriente trifásica de la subestación y de todas las fuentes de GD. Ángulo de falla de las corrientes.	Menores al 1%	[52]
MLPNN y RBFNN	2	1. Cuatro tipos diferentes de fallas a 120 metros de longitud de la línea 1. 2. Cuatro tipos diferentes de fallas a 360 metros de longitud de la línea 13. 3. Cuatro tipos diferentes de fallas a 850 metros de longitud de la línea 18. 4. Cuatro tipos diferentes de fallas a 570 metros de longitud de la línea 21.	22 barras, 21 líneas, 20 kV y 22 cargas.	No indica.	Corrientes trifásicas de la subestación y de todas las fuentes de GD	Máximo error de 8.33% para RBFNN. Máximo error de 17.5% para MLPNN	[53]
MLPNN	2	1. Falla trifásica	22 barras, 21 líneas, 20 kV y 22 cargas.	0,50,100 y 150.	Corrientes trifásicas de la subestación y de todas las fuentes de GD	1.4%	[12]
		2. Falla monofásica				0.3%	
		3. Falla bifásica				0.018%	
		4. Falla bifásica a tierra				0.13%	

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
RBFNN	6	1. Falla monofásica a tierra a 380 metros de longitud de la línea 1. 2. Falla bifásica a 430 metros de longitud de la línea 3. 3. Falla bifásica a tierra a 680 metros de longitud de la línea 10. 4. Falla trifásica a 870 metros de longitud de la línea 19.	32 barras, 31 líneas, 20 kV y 32 cargas.	No indica.	Corrientes trifásicas de la subestación y de todas las fuentes de GD	1. 2.3% 2. 0.4% 3. 0.24% 4. 0.14%	[54]
RB-FNN-OSD					Corrientes trifásicas de la subestación y de todas las fuentes de GD	1. 0% 2. 0% 3. 0.02% 4. 0.014%	

Tabla 9. Resultados de aplicación de las SVM en la localización de fallas en SED con GD

Método	Número de fuentes de GD	Consideración	Circuito bajo estudio	Resistencia de falla (Ω)	Datos de entrada	Error en la localización	Fuente
SVM-C	1	Simulación de los diferentes tipos de falla en todas las líneas del circuito. Se determina la línea en donde ocurre la falla Cantidad de patrones de entrenamiento: 10500 Cantidad de patrones de prueba: 2268	22 barras, 11 kV, 21 líneas	0, 25, 30, 50, 60, 75, 90 y 100.	Valores eficaces de las tensiones y corrientes trifásicas medidas en la subestación y en todas las fuentes de GD	0%	[13]
	2					0.05%	
	3					7.94%	
SVM-R	1	Se determina la localización de la falla como porcentaje de longitud de la línea donde ocurre. Se determina el valor de la resistencia de falla. Cantidad de patrones de entrenamiento: 10500. Cantidad de patrones de prueba: 2268				No indica	
	2					Máximo error 0.8%	
	3					Máximo error 0.9%	
SVM-C	1	Simulación de los diferentes tipos de falla en todos los nodos del circuito. Se determina la zona donde ocurre la falla. Se utilizan dos (2) medidores inteligentes. Cantidad de patrones de entrenamiento: 1235 Cantidad de patrones de prueba: 529	IEEE de 34 barras, 24.9 kV y una parte en 4.16 kV, 33 lineas	5, 10, 15, 20, 25 y 30.	Valores eficaces de las tensiones trifásicas medidas en la subestación y en la fuente de GD	Máximo error de 12% para falla monofásica y mínimo error de 0% para falla trifásica	[45]

Tabla 10. Aplicación de métodos integrados para localizar fallas en SED

Métodos utilizados	Consideración	Observación	Fuente
Convencionales IA	La falla se detecta y clasifica utilizando la transformada de wavelet. Se emplea un método basado en la impedancia para identificar todas las posibles localizaciones de fallas, y una ANN estima la sección en falla para identificar la ubicación más probable.	No considera GD	[55]
Convencionales IA	Utiliza una combinación de una técnica de clasificación y un método basado en la impedancia para estimar el tipo y la ubicación de la falla. El clasificador se basa en un algoritmo de IA para encontrar la zona en falla e identificar la ubicación correcta entre diferentes localizaciones de fallas estimadas.	No considera GD	[56]
Convencionales	Los autores integran un método basado en la impedancia con un método basado en la onda itinerante para superar el problema de las estimaciones múltiples. Primero utilizan un algoritmo basado en la impedancia para estimar la distancia de falla y encontrar todas las localizaciones posibles. A continuación, se calculan las frecuencias características asociadas con cada localización de falla posible y se realiza un análisis transitorio para identificar los componentes de frecuencia más significativos de las ondas viajeras generadas por fallas. Por último, se realiza un análisis de correlación tanto en el dominio de la frecuencia como en el dominio del tiempo para encontrar la solución correcta.	No considera GD	[57]
Convencionales	En primer lugar, se propone un algoritmo basado en la impedancia para encontrar todas las posibles ubicaciones de la falla y luego se proponen dos métodos para determinar la localización real de la falla. Ambos métodos simulan el mismo tipo de falla en cada localización posible de falla y registran la tensión al principio del alimentador. Para encontrar la localización real de la falla, el primer método compara las muestras de tensión medidas y simuladas, mientras que el segundo método coincide con el espectro de frecuencias de transitorios de fallas registradas y simuladas.	No considera GD	[58]
Convencionales	Propone el uso de medidores inteligentes con capacidad de monitoreo de tensión para encontrar el área donde la magnitud de tensión es baja debido a la proximidad de la falla. El método propuesto construye una zona de baja tensión para identificar la ubicación real de la falla y mejorar la eficiencia de los métodos basados en la impedancia.	No considera GD	[59]
Convencionales IA	Se diseña una estrategia inteligente como contribución a la exactitud en la localización de fallas en SED aéreos con GD. Se genera una base compuesta por 1764 datos de falla que consideran los 11 tipos diferentes de fallas que se pueden presentar en un SED para seis (6) valores diferentes de resistencia de falla que van de cinco (5) a 30 Ohm, en pasos de 5 Ohm. La estrategia es integrada al hacer uso de la inteligencia artificial y de algoritmos matemáticos basados en el modelo del sistema, se propone por medio de tres (3) etapas: Identificación del tipo de falla, localización de la zona de falla y localización del punto exacto de falla.	Considera la presencia de GD	[60]

Tabla 11. Comparación de diferentes métodos propuestos de localización de fallas en SED con GD

Método	Tipo	Ventajas	Limitaciones	Fuente
Jamali y Talavat	Basado en la impedancia	No requiere la medición en la fuente de GD	Solo para máquina síncrona como fuente de GD	[61]
Orozco *et al.*	Basado en la impedancia	No requiere los fasores de tensión sincronizada en la fuente de GD	El método es sensible a errores de medición e inexactitudes de datos de carga	[62]
Alwash *et al.*	Basado en la impedancia	Supera el requisito de identificación de tipo de falla	El análisis del efecto de los errores de medición puede ayudar a probar el método	[29]
Brahma	Basado en mediciones dispersas	Algorítmicamente simple	La estimación inexacta de la corriente de falla afecta la exactitud de los resultados, especialmente en el caso de bajas resistencias de falla	[30]
Chen *et al.*	Basado en mediciones dispersas	Algorítmicamente simple	Sensible a las inexactitudes de los datos de carga. Complicado para fallas trifásicas y fallas doble línea a tierra.	[31]
Javadian *et al.*	Inteligencia artificial	Tiempo de ejecución muy corto	Requiere datos de entrenamiento y re-entrenamiento por cualquier cambio en la topología del SED	[63]
Zayandehroodi *et al.*	Inteligencia artificial	Tiempo de ejecución muy corto	Requiere datos de entrenamiento y re-entrenamiento por cualquier cambio en la topología del SED	[54]

Método	Tipo	Ventajas	Limitaciones	Fuente
Rafinia y Moshtagh	Inteligencia artificial	Tiempo de ejecución muy corto. Clasifica y localiza todos los tipos de falla serie y paralelo	Requiere datos de entrenamiento y re-entrenamiento por cualquier cambio en la topología del SED. Requiere mediciones con frecuencia de muestreo comparativamente altas	[64]
Pérez y Vásquez	Inteligencia Artificial	Tiempo de ejecución muy corto. Localiza todos los tipos de falla paralelo que se pueden presentar en los SED. Considera diferentes valores de resistencia de falla.	Requiere datos de entrenamiento y re-entrenamiento por cualquier cambio en la topología del SED. Solo localiza la zona donde ocurre la falla sin determinar el punto exacto de ubicación.	[45]
Brahma y Girgis	Otros (compara la contribución de la corriente de falla de cada fuente con las corrientes previamente calculadas mediante análisis de cortocircuito). Algoritmos utilizando indicadores de fallos	Algorítmicamente simple	Sensible a la resistencia de falla. Requiere una gran cantidad de análisis de cortocircuito	[65]
Conti y Nicotra	Algoritmos utilizando indicadores de fallos	Algorítmicamente simple	Requiere dispositivos de detección de fallas en todas las subestaciones secundarias	[66]
Jamali y Talavat	Otros (basados en un análisis numérico de dominio de tiempo usando un modelo de máquina síncrono de orden completo)	No requiere la medición en la fuente de GD	Solo para máquina síncrona como fuente de GD	[67]
Jamali *et al.*	Otros (utiliza algoritmo genético y medidas dispersas)	En comparación con (P. Chen *et al.*, 2015; Lotfifard, Kezunovic, y Mousavi, 2011), pasa por alto la estimación de impedancia de falla y busca un número limitado de nodos para encontrar la falla	Solo fue probado para SED con un número limitado de nodos	[69]

Referencias

[1] T. Short, *Electric power distribution: handbook*, vol. 10. 2004.

[2] S. Ramirez, *Redes de Distribución de Energía*, 3.a ed. Manizales, 2004.

[3] M. Kezunovic, «Smart Fault Location for Smart Grids», *IEEE Transactions on Smart Grid*, vol. 2, n.o 1, pp. 11-22, 2011.

[4] P. Mohammadi, E. Kishyky, A. Akher, y A. Salam, «The impacts of distributed generation on fault detection and voltage profile in power distribution networks», *International Power Modulator and High Voltage Conference (IPMHVC)*, pp. 191-196, 2014.

[5] Y. Menchafou, H. Markhi, M. Zahri, y M. Habibi, «Impact of distributed generation integration in electric power distribution systems on fault location methods», *3rd International Renewable and Sustainable Energy Conference (IRSEC)*, n.o 1998, 2015.

[6] A. Fazanehrafat, S. Javadian, S. Bathaee, y M. Haghifam, «Maintaining the recloser-fuse coordination in distribution systems in presence of DG by determining DG's size», *IET 9th International Conference on Developments in Power Systems Protection (DPSP 2008)*, pp. 132-137, 2008.

[7] B. José, P. Cavalcante, F. Trindade, y M. Almeida, «Analysis of Distance Based Fault Location Methods for Smart Grids with Distributed Generation», *2013 4th IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1-6, 2013.

[8] C. Orozco, J. Mora, y S. Pérez, «Método de localización de fallas basado en impedancia aparente para sistemas de distribución con generación distribuida», *Ingeniare Revista chilena de ingeniería*, vol. 23, n.o 3, pp. 348-360, 2015.

[9] J. Mora, J. Meléndez, y G. Carrillo, «Comparison of impedance based fault location methods for power distribution systems», *Electric Power Systems Research*, vol. 78, n.o 4, pp. 657-666, 2008.

[10] L. Awalin, H. Mokhlis, y A. Bakar, «Recent developments in fault location methods for distribution networks», *Przeglad Elektrotechniczny*, n.o 12, pp. 206-212, 2012.

[11] M. Mirzaei, M. Kadir, E. Moazami, y H. Hizam, «Review of Fault Location Methods for Distribution Power System», *Australian Journal of Basic and Applied Sciences*, vol. 3, n.o 3, pp. 2670-2676, 2009.

[12] S. Javadian y M. Massaeli, «A fault location method in distribution networks including DG», *Indian Journal of Science and Technology*, vol. 4, n.o 11, pp. 1446-1451, 2011.

[13] R. Agrawal y D. Thukaram, «Identification of fault location in power distribution system with distributed generation using support vector machines», *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1-6, 2013.

[14] M. Torres, «Contribución de la información en el diseño de una aproximación de la evaluación de la calidad del servicio», *Trabajo de Tesis Doctoral*

presentada para optar al grado de Doctor en Ciencias de la Ingeniería mención Productividad, Universidad Experimental Politécnica" Antonio José de Sucre", UNEXPO Barquisimeto, 2010.

[15] R. Das, «Determining the locations of faults», University of Saskatchewan, 1998.

[16] C. Vásquez, «Criterios de eficiencia para disminuir las pérdidas debidas a la no calidad de la energía eléctrica», en *Trabajo de Ascenso presentado ante la Universidad Nacional Experimental Politécnica «Antonio José de Sucre» como requisito para ascender a la Categoría de Profesora Titular*, Barquisimeto, 2011, pp. 1-164.

[17] IEEE Power & Energy Society, «IEEE Recommended Practice for Monitoring Electric Power Quality», *IEEE Std 1159*, 2009.

[18] J. Burke y D. Laurence, «Characteristics of fault currents on distribution systems», *IEEE Transactions on Power Apparatus and Systems*, n.o 1, pp. 1-6, 1984.

[19] H. Farooq, C. Zhou, M. E. Farrag, y M. Ejaz, «Investigating the impacts of distributed generation on an electrical distribution system already stressed by non-linear domestic loads», *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, vol. 100, n.o 1, pp. 2-5, 2012.

[20] T. Ackermann, G. Andersson, y L. Söder, «Distributed generation: A definition», *Electric Power Systems Research*, vol. 57, n.o 3, pp. 195-204, 2001.

[21] S. Gururajapathy, H. Mokhlis, y H. Illias, «Fault location and detection techniques in power distribution systems with distributed generation: A review», *Renewable and Sustainable Energy Reviews*, vol. 74, n.o February 2016, pp. 949-958, 2017.

[22] R. Acevedo, «Análisis de la evolución del marco legal del servicio eléctrico venezolano en el período 2000 al 2010», *Gestión y Gerencia*, vol. 5, n.o 3, pp. 104-122, 2010.

[23] J. Montenegro y A. Urdaneta, «Evolución histórica del marco regulatorio del sector eléctrico latinoamericano», *Interciencia*, vol. 39, n.o 3, pp. 164-171, 2014.

[24] IEEE Power Engineering Society, «IEEE Guide for Electric Power Distribution Reliability Indices», *IEEE Std 1366*, pp. 1-21, 1998.

[25] A. Bahmanyar, S. Jamali, A. Estebsari, y E. Bompard, «A comparison framework for distribution system outage and fault location methods», *Electric Power Systems Research*, vol. 145, pp. 19-34, 2017.

[26] D. Novosel, D. Hart, Y. Hu, y J. Myllymaki, «System for locating faults and estimating fault resistance in distribution networks with tapped loads», 1998.

[27] M. Saha y E. Rosolowski, «Method and device of fault location for distribution networks,US Patent number 6,483,435», 2002.

[28] J. Zhu, D. Lubkeman, y A. Girgis, «Automated fault location and diagnosis on electric power distribution feeders», *Power Delivery, IEEE Transactions*, vol. 12, n.o 2, pp. 801-809, 1997.

[29] S. Alwash, V. Ramachandaramurthy, y N. Mithulananthan, «Fault-Location Scheme for Power Distribution System with Distributed Generation», *IEEE Transactions on Power Delivery*, vol. 30, pp. 1187-1195, 2015.

[30] S. Brahma, «Fault location in power distribution system with penetration of distributed generation», *IEEE Transactions on Power Delivery*, vol. 26, n.o 3, pp. 1545-1553, 2011.

[31] P. Chen, V. Malbasa, Y. Dong, y M. Kezunovic, «Sensitivity Analysis of Voltage Sag Based Fault Location With Distributed Generation», *IEEE Transactions on Smart Grid*, vol. 6, n.o 4, pp. 2098-2106, 2015.

[32] A. Bretas y R. Salim, «Fault location in unbalanced dg systems using the positive sequence apparent impedance», *2006 IEEE PES Transmission and Distribution Conference and Exposition: Latin America, TDC'06*, vol. 00, pp. 1-6, 2006.

[33] J. Nunes y A. Bretas, «A impedance-based fault location technique for unbalanced distributed generation systems», *2011 IEEE PES Trondheim PowerTech: The Power of Technology for a Sustainable Society, POWERTECH 2011*, pp. 1-7, 2011.

[34] J. Nunes y A. Bretas, «Impedance-based fault location formulation for unbalanced primary distribution systems with distributed generation», *2010 International Conference on Power System Technology: Technological Innovations Making Power Grid Smarter, POWERCON2010*, 2010.

[35] D. Penkov, B. Raison, C. Andrieu, J. P. Rognon, y B. Enacheanu., «DG impact on three fault location. DG use for fault location purposes?», *International Conference on Future Power Systems*, vol. 4, n.o December, p. 6, 2005.

[36] J. Cordova y O. Faruque, «Fault location identification in smart distribution networks with distributed generation», 2015.

[37] F. Dehghani y H. Nezami, «A new fault location technique on radial distribution systems using artificial neural network», *22nd International Conference on Electricity Distribution*, pp. 10-13, 2013.

[38] A. Abur, «Fault Location in Distribution Systems with Distributed Generation», en *15th PSCC*, 2005, pp. 1-5.

[39] X. Chen, X. Yin, y S. Deng, «A novel method for SLG fault location in power distribution system using time lag of travelling wave components», *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 12, n.o 1, pp. 45-54, 2017.

[40] X. Chen, X. Yin, X. Yin, J. Tang, y M. Wen, «A novel traveling wave based fault location scheme for power distribution grids with distributed generations», *2015 IEEE Power & Energy Society General Meeting*, pp. 1-5, 2015.

[41] M. Saha, J. Izykowski, y E. Rosolowski, *Fault Location on Power Networks*, 1.a ed. Londrés, 2010.

[42] C. Cortes y V. Vapnik, «Support-Vector Networks», *Machine Learning*, vol. 20, n.o 3, pp. 273-297, 1995.

[43] S. R. Gunn, «Support Vector Machines for Classification and Regression», *Image Speech and Intelligent Systems Technical Report*, vol. 14, n.o May, pp. 230-67, 1998.
[44] V. Vapnik, S. E. Golowich, y A. Smola, «Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing», *Annual Conference on Neural Information Processing Systems (NIPS)*, pp. 281-287, 1996.
[45] R. Pérez y C. Vásquez, «Fault Location in Distribution Systems with Distributed Generation Using Support Vector Machines and Smart Meters», *IEEE Ecuador Technical Chapters Meeting (ETCM)*, pp. 1-6, 2016.
[46] B. E. Boser, I. M. Guyon, y V. N. Vapnik, «A Training Algorithm for Optimal Margin Classifiers», *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, pp. 144-152, 1992.
[47] C. Hsu y C. Lin, «A comparison of methods for multiclass support vector machines», *IEEE Transactions on Neural Networks*, vol. 13, n.o 2, pp. 415-425, 2002.
[48] M. Daisy y R. Dashti, «Single phase fault location in electrical distribution feeder using hybrid method», *Energy*, vol. 103, pp. 356-368, 2016.
[49] M. Hagh, M. Hosseini, y S. Asgarifar, «A novel phase to phase fault location algorithm for distribution network with distributed generation», *CIRED Workshop*, pp. 1-4, 2012.
[50] F. Abo, M. Alaam, y A. Azmy, «Impedance-based fault location technique for distribution systems in presence of distributed generation», *IEEE International Conference on Smart Energy Grid Engineering, SEGE 2013*, pp. 28-30, 2013.
[51] A. Bretas, M. Moreto, R. Salim, y L. Pires, «A Novel High Impedance Fault Location for Distribution Systems Considering Distributed Generation», *2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America*, n.o September, pp. 1-6, 2006.
[52] G. Shahgholian y M. Rezaei, «Fault Location Scheme in Distribution Systems with Distributed Generators Using Neural Networks», *Majlesi Journal of Electrical Engineering*, vol. 4, n.o 2, pp. 57-62, 2010.
[53] H. Zayandehroodi, A. Mohamed, H. Shareef, y M. Mohammadjafari, «Performance comparison of mlp and rbf neural networks for fault location in distribution networks with DGs», *2010 IEEE International Conference on Power and Energy*, pp. 341-345, 2010.
[54] H. Zayandehroodi, A. Mohamed, M. Farhoodnea, y M. Mohammadjafari, «An optimal radial basis function neural network for fault location in a distribution network with high penetration of DG units», *Measurement: Journal of the International Measurement Confederation*, vol. 46, n.o 9, pp. 3319-3327, 2013.
[55] R. Salim, K. de Oliveira, A. Filomena, M. Resener, y A. Bretas, «Hybrid fault diagnosis scheme implementation for power distribution systems automa-

tion», *IEEE Transactions on Power Delivery*, vol. 23, n.o 4, pp. 1846-1856, 2008.
[56] J. Mora, V. Barrera, y G. Carrillo, «Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis», *IEEE Transactions on Power Delivery*, vol. 22, n.o 3, pp. 1715-1721, 2007.
[57] D. Gazzana, G. Ferreira, A. Bretas, A. Bettiol, A. Carniato, L. Passos, A. Ferreira, y J. Silva, «An integrated technique for fault location and section identification in distribution systems», *Electric Power Systems Research*, vol. 115, pp. 65-73, 2014.
[58] R. Dashti y J. Sadeh, «Fault section estimation in power distribution network using impedance-based fault distance calculation and frequency spectrum analysis», *IET Generation, Transmission & Distribution*, vol. 8, n.o 8, pp. 1406-1417, 2014.
[59] F. Trindade y W. Freitas, «Low Voltage Zones to Support Fault Location in Distribution Systems With Smart Meters», *IEEE Transactions on Smart Grid*, vol. pp, n.o 99, pp. 1-10, 2016.
[60] R. Pérez, «Estrategia inteligente como contribución a la exactitud en la localización de fallas en sistemas eléctricos de distribución aéreos con generación distribuida», UNEXPO, 2017.
[61] S. Jamali y V. Talavat, «Accurate fault location method in distribution networks containing distributed generations», *Iranian Journal of Electrical and Computer Engineering*, vol. 10, n.o 1, pp. 27-33, 2011.
[62] C. Orozco, A. Bretas, R. Chouhy, A. Herrera, y J. Marín, «Active distribution network fault location methodology: A minimum fault reactance and Fibonacci search approach», *International Journal of Electrical Power and Energy Systems*, vol. 84, pp. 232-241, 2017.
[63] S. Javadian, M. Haghifam, y N. Rezaei, «A fault location and protection scheme for distribution systems in presence of DG using MLP neural networks», *2009 IEEE Power and Energy Society General Meeting, PES '09*, pp. 1-8, 2009.
[64] A. Rafinia y J. Moshtagh, «A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ANN and FLS», *International Journal of Electrical Power and Energy Systems*, vol. 55, pp. 261-274, 2014.
[65] S. Brahma y A. Girgis, «Development of adaptive protection scheme for distribution systems with high penetration of distributed generation», *2003 IEEE Power Engineering Society General Meeting*, vol. 4, n.o 1, pp. 56-63, 2003.
[66] S. Conti y S. Nicotra, «Procedures for fault location and isolation to solve protection selectivity problems in MV distribution networks with dispersed generation», *Electric Power Systems Research*, vol. 79, n.o 1, pp. 57-64, 2009.
[67] S. Jamali y V. Talavat, «Dynamic fault location method for distribution networks with distributed generation», *Electrical Engineering*, vol. 92, n.o 3, pp. 119-127, 2010.

[68] S. Lotfifard, M. Kezunovic, y M. Mousavi, «Voltage sag data utilization for distribution fault location», *IEEE Transactions on Power Delivery*, vol. 26, n.o 2, pp. 1239-1246, 2011.
[69] S. Jamali y A. Bahmanyar, «A Fast and Accurate Fault Location Method for Distribution Networks with DG Using Genetic Algorithms», en *Smart Grid Conference (SGC), 2015*, 2015, pp. 110-114.

Sobre los autores



Ramón Enrique Pérez Pineda

Graduado en el año 2006 como Ingeniero Electricista en la Universidad Nacional Politécnica Antonio José de Sucre (UNEXPO) en Barquisimeto, Venezuela. En los años 2011 y 2018 la UNEXPO le otorga los títulos de Magister Scientiarum en Ingeniería Eléctrica y Doctor en Ciencias de la Ingeniería. Ha sido autor y coautor de varios trabajos de investigación indexados en diferentes bases de datos. Sus principales áreas de investigación son las máquinas eléctricas, sistemas eléctricos de distribución y microredes eléctricas. Actualmente es profesor de la Universidad Politécnica Salesiana en Quito, Ecuador y director de las carreras de Ingeniería Electrónica - Electrónica y Automatización, además de coordinar el Grupo de Investigación en Electrónica, Control y Automatización (GIECA).



Carmen Luisa Vásquez Stanescu

De nacionalidad venezolana y rumana. Ingeniero Electricista graduada de la Universidad Nacional Politécnica Antonio José de Sucre (UNEXPO) en 1995. Magister Scientiarum en Ingeniería Eléctrica en 1998 en la misma institución. Doctora en Ciencias Técnicas del Instituto Superior Politécnico "José Antonio Echeverría" de La Habana, Cuba, en el año 2002. Desde 1995 se ha desempeñado como docente e investigadora de la UNEXPO Vicerrectorado Barquisimeto, en las áreas de investigación de Calidad de la Energía Eléctrica, Ingeniería de la Calidad, Transitorios Electromagnéticos y Eficiencia Energética. Adicionalmente, es Coordinadora de la Maestría de Ingeniería Eléctrica, miembro del Comité Académico del Doctorado en Ciencias de la Ingeniería y, desde el año 2014, es Coordinadora de la Unidad Central de Investigación de la UNEXPO. Ha participado en la Acción de Coordinación EFESOS del CYTED (2007-2009) y, actualmente, la Red RITMUS (2018-2021). Ha promovido y participación en 15 proyectos de investigación con financiamiento nacional e internacional, con instituciones AECID, CYTED, Istituto Nazionale Di Fisica Nucleare de Legnaro National Laboratories (Italia), Centro para la Cooperación y para el Desarrollo (CDD) y CITCEA de la Universidad Politécnica de Cataluña (UPC), Barcelona, España, entre otros. Autora del Libro "Eficiencia Energética y calidad de la energía eléctrica. Criterios de diseño, operación y mantenimiento" (2010) y coautora del libro "Eficiencia Energética. Aplicaciones a la

seguridad del suministro y la sostenibilidad el medio ambiente” (2011). Autora y ponente de 140 trabajos de investigación indexados en diferentes bases de datos.



Amelec Jesús Viloria Silva
Ingeniero Químico con doctorado en Ciencias de la Ingeniería, egresado de la Universidad Nacional Experimental Politécnica Antonio José de Sucre en Barquisimeto, Venezuela. Más de 200 artículos internacionales avalan su actividad investigativa en el área de la estadística y la gestión de productividad. Actualmente es Profesor titular en la Universidad de la Costa, Barranquilla Colombia.

CAPÍTULO III

Evaluación de la cosecha de neblina como una fuente potencial para el aprovechamiento de agua

Evaluation of fog water collection as a potential source of water supply

Christian Domínguez*, Paola Echeverría§, Marcos Villacís†, Sophie Violette†
*Universidad Politécnica Salesiana Ecuador, §†Escuela Politécnica Nacional Ecuador, †Universidad Sorbona Francia

*email: cdominguez@ups.edu.ec
§email: pao_echev1@hotmail.com
†email: marcos.villacis@epn.edu.ec
†email:sophie.violette@ens.fr

Resumen

El término “ciudades inteligentes” es asociado únicamente con el empleo de tecnologías de la información en la infraestructura de una ciudad. Sin embargo, un aspecto fundamental para satisfacer las necesidades de una población es el aprovechamiento de recursos naturales, entre ellos el agua. En este sentido, las técnicas de aprovechamiento de agua deben ser sostenibles y amigables con el ambiente, este puede ser el caso de la cosecha de neblina. En este estudio, se presenta una metodología para cuantificar la intercepción de neblina por un sistema de captación y evaluar su potencial como una fuente de agua alternativa. Esta metodología fue probada en un sitio experimental en la parte alta de la isla San Cristóbal (Galápagos). Se realizó un monitoreo de dos meses que incluyó la medición de las principales variables meteorológicas y la intercepción de neblina. Se utilizaron tres colectores de neblina, un cilíndrico (CFC), y dos colectores estándar de 35% (SFC35) y 50% (SFC50) de coeficiente de sombra, respectivamente. Dado que los colectores captan tanto neblina como lluvia, se utilizó un modelo geométrico para separar estas dos variables. Los resultados de este estudio muestran que los colectores CFC, SFC35 y SFC50 tienen un rendimiento de 3.7, 4.1 y 4.8 mm/día, respectivamente. El colector SFC50 es el que tiene un mejor rendimiento de intercepción de neblina,

el que puede ser atribuido a su alta eficiencia aerodinámica comparada con los otros. Al analizar las variables meteorológicas se evidencia que el potencial de captura de neblina puede ser aún mayor.

Palabras clave

Ciudades inteligentes, clima, galápagos, intercepción de neblina, sostenibilidad

Abstract

The term "smart cities" is usually associate to the use of information technologies in a city infrastructure. Nevertheless, the use of natural resources, such as water, is an important aspect to satisfy the inhabitants necessities. In this sense, techniques to supply water must be sustainable and environmentally friendly, which is the case of fog water collection. In this study, we present a methodology to quantify fog water intercepted by a fog gauge system and to assess its potential as a source of water supply. This methodology is tested in an experimental site located at the highlands of San Cristobal Island (Galapagos), which included two months monitoring of the main meteorological variables and fog interception. Three fog gauges were used to quantify fog interception, a cylindrical (CFC) and two standards of 35% (SFC35) and 50% (SFC50) shading coefficient, respectively. Given that fog gauge collectors capture both fog and rainfall, we used a geometrical model to separate both variables. Results show that collectors' CFC, SFC35 and SFC50 have an average performance of 3.7, 4.1 and 4.8 mm/day, respectively. SFC50 collector presents the best performance which can be related to its relative higher aerodynamic efficiency. The analysis of the meteorological variables evidences that fog water collection potential can be even higher at this site.

Keywords

Smart cities, climate, Galapagos, fog water interception, sustainability

1. Introducción

En los últimos años, el término "ciudades inteligentes" se ha utilizado con mayor frecuencia en la política de gobiernos y municipios, en artículos científicos, e inclusive en medios de comunicación masivos y redes sociales. Esta popularidad se alcanzó, debido al constante debate de cómo las nuevas tecnologías podrían ayudar a mejorar la habitabilidad en áreas urbanas.

No existe un concepto claro sobre qué son las ciudades inteligentes, debido a que su definición depende de la visión que se tenga de las mismas [1]. Mientras varios autores relacionan este concepto con el empleo de tecnologías de la información en la infraestructura

de una ciudad o al desarrollo de negocios (ej. [2]–[7] otros autores orientan este concepto hacia el bienestar y desarrollo de sus habitantes (ej. [8]–[11]) , y pocos hacia el tema ambiental (ej. [12]–[14]), como el manejo adecuado de los recursos naturales. Tanto Albino *et al*., 2015 [1] y Batisha, 2015 [15] concluyen que dicho concepto es único para cada ciudad y depende de las necesidades de la misma.

Sin duda alguna, una de las necesidades más importantes de toda ciudad, es el aprovechamiento de recursos naturales, entre ellos el agua. Además de que el agua es un elemento fundamental para la supervivencia de los seres vivos, también es un factor importante en los procesos de producción, para el desarrollo económico de una población e inclusive para la recreación de sus habitantes. Sin embargo, muchas regiones en el mundo sufren de falta de agua [16]–[18]. En al año 2000, de 1.2 a 2.4 billones de personas (20 al 39% de la población mundial) vivía en cuencas hidrográficas expuestas a falta de recursos hídricos [19]–[24]. Tanto el incremento de la población y el crecimiento del desarrollo económico global, son factores importantes para el crecimiento de la escasez del agua de los habitantes. También, el cambio climático constituye un factor clave para la evaluación de la disponibilidad de recursos hídricos, cambios en la temperatura y en los patrones de precipitación resultan una amenaza para el abastecimiento de agua en varios sitios. Varios autores [16], [18]–[25] han utilizado diferentes modelos hidrológicos y modelos climáticos globales con diversos posibles escenarios para estimar la disponibilidad de agua en años futuros, se espera que para el año 2050, entre el 37 y 53% de población global viva en regiones con problemas de falta de agua. En la actualidad, muchos países ya están tomando conciencia de este asunto. Países como Australia [26], China [27], España [28], [29] y Estados Unidos [30], ya están contemplando medidas y se encuentran buscando alternativas para asegurar el abastecimiento de sus poblaciones. En este contexto, una ciudad inteligente debe tomar en cuenta el problema de fuentes de agua para el suministro de sus habitantes.

En la actualidad, el abastecimiento de agua para las poblaciones proviene de dos fuentes principales: i) el agua superficial de ríos y lagunas, conocida como “agua azul”, en donde el agua es tomada a través de infraestructura hidráulica; y ii) el agua subterránea,

conocida como “agua verde”, en donde el agua es extraída de acuíferos mediante el bombeo de pozos y usos de galerías. Es claro que las poblaciones que se ubican cerca de fuentes superficiales de agua pueden aprovechar este recurso. Sin embargo, no se tiene certeza sobre el efecto que tendría el cambio climática en la disponibilidad de las misma [18], dicha incertidumbre podría contemplar una disminución en su cantidad. Las poblaciones que se benefician de aguas subterráneas tienen la dificultad de que estas fuentes son muy susceptibles a la contaminación o a la intrusión del agua del mar (en el caso de zonas costeras), además de la incertidumbre en cuanto al cambio climático (existe la posibilidad que las precipitaciones sean más intensas y esporádicas, generando mayor escorrentía y posiblemente reduciendo la recarga de acuíferos). Existen otras fuentes de agua alternativa como: el reciclamiento de aguas residuales, la desalinización de agua del mar o el aprovechamiento de la neblina, este último conocido como “cosecha de neblina”. Las dos primeras alternativas, aunque resulten en fuentes prácticamente ilimitadas de agua, su potabilización requiere de una cantidad considerable de energía, y actualmente se utiliza combustibles fósiles en su mayoría [28]; motivo por el cual estos sistemas no podrían ser considerados como sustentables. La cosecha de neblina, podría resultar en una fuente de agua sustentable, el equipamiento y mantenimiento es sencillo y relativamente barato. Este sistema se lo ha utilizado en pequeñas comunidades con resultados satisfactorios [15], [31], [32]. El problema de la cosecha de agua es que está limitado a zonas expuestas a neblina y que la cantidad potencial de agua recolectada por este sistema es desconocida hasta cierto punto.

La neblina consiste en pequeñas gotas de agua que se encuentran en suspensión debido a que su tamaño no es lo suficientemente grande para que su velocidad de caída supere la velocidad de empuje del aire [33], permitiéndolas precipitarse. La neblina se puede formar por diversos procesos físicos en la atmósfera [34], [35], pero en general ocurre debido a la condensación de la evaporación de la humedad del suelo, vegetación, de masas de agua o de masas de aire húmedas. Cuando la neblina es llevada por el viento, las pequeñas gotas de agua chocan con la estructura de la vegetación, estas coalescen formando gotas más grandes que finalmente se precipitan

en el suelo [33], [34], [36]. A este proceso se conoce como intercepción de neblina, y puede ser una fuente importante de agua para varios ecosistemas, alcanzando valores entre 4 y 98% del aporte de la precipitación [37]–[43].

Los habitantes de regiones con escasez de fuentes de agua convencionales han tratado de aprovechar este proceso hidrológico como una fuente de agua alternativa, al utilizar colectores para captar neblina. Existen varios proyectos de sistemas de colección de neblina alrededor del mundo (Chile, Ecuador, Sudáfrica, Namibia, Omán, Arabia Saudita, España). De acuerdo a la literatura, la producción de diferentes colectores de neblina oscila entre 2 y 10 l/m^2día en promedio [15], [31], [51], [52], [32], [44]–[50], alcanzando valores de hasta 300 $l/m^2 \cdot$día [45]. De acuerdo a Batisha, 2007 [15], la implementación de estos sistemas puede resultar atrayente debido a que se puede colectar una gran cantidad de agua, su costo es relativamente bajo o comparable con otros sistemas de agua, los colectores son simples y no requieren energía externa, y porque el agua captada se espera que sea de buena calidad. Sin embargo, la factibilidad de este sistema depende de un factor importante como es la cantidad de neblina captable. Antes de la implementación de este tipo de sistema se debe realizar una investigación sobre la cuantificación de neblina en el sitio con diferentes colectores de neblina y que incluya el monitoreo de otras variables meteorológicas que permitan conocer las condiciones que favorecen la intercepción de neblina. Debido a que estos sistemas captan tanto lluvia como neblina, sería también importante separar estas componentes para evaluar realmente el aporte de la neblina.

Conociendo la problemática del incremento de la demanda de agua, uno de los componentes que se tiene que trabajar alrededor de una ciudad inteligente es el manejo y aprovechamiento de los recursos hídricos. Sin duda alguna, los sistemas de aprovechamiento de agua deberían ser sustentables, con tecnologías que sean amigables con el ambiente y que incluya la participación de la población. Uno de estos sistemas puede ser la captación de neblina. En este trabajo presentamos una metodología simple para evaluar la cantidad de neblina recolectada por un sistema de captación de neblina, determinando no solamente la cantidad de neblina captada sino también identificando

los factores que afectan su captación. Este trabajo se estructura de la siguiente manera: primero, se presenta la metodología y equipamiento para realizar esta evaluación, luego, se propone un caso de estudio en la isla San Cristóbal (Galápagos), expuesta a la presencia de neblina y con problemas de escasez de agua para su sostenibilidad. A continuación, los resultados del caso de estudio son presentados y discutidos. Finalmente las conclusiones del trabajo son presentadas.

2. Metodología

A. Intercepción de neblina y ubicación geográfica

Existen básicamente tres factores que afectan la intercepción de neblina [33], [53], [54] la ubicación geográfica, diseño del colector y factores climáticos. Mientras las gotas de lluvia tienen un diámetro de 0.5 a 5 mm, las gotas de agua en la neblina tiene un diámetro de 1 um a 40 um [32], [35]. Este pequeño tamaño es el motivo por el cual estas gotas no pueden precipitarse, y por ende flotan sobre la superficie del suelo como niebla, neblina o bruma. Del punto de vista meteorológico, la diferencia entre estos tres últimos nombres radica básicamente en el rango de visibilidad que permiten (ej. la variable climática que reduce la visibilidad a una distancia menor a 1000 m se considera como niebla). Para el punto práctico de este estudio, dicha diferenciación no es relevante, y por lo tanto a la niebla y bruma se los considerará como neblina. La intercepción de neblina ocurre cuando las pequeñas gotas de agua que conforman la neblina son acarreadas por el viento y chocan contra una superficie, que se satura y permite su precipitación.

La ubicación geográfica es un factor clave para la implementación de un sistema de cosecha de neblina, ya que no todas las regiones están expuestas a la presencia de neblina. Las regiones costeras serían las mayormente favorecidas por este sistema, ya que están propensas a la neblina de advección, producto de la condensación de las masas de aire caliente y cargadas de humedad, provenientes del mar, cuando pasan por el suelo frío. Este tipo de neblina puede estar presente varios cientos de kms hacia la superficie continental, es por este motivo que varias regiones áridas (ej. desierto de Ataca-

ma, desierto del Sahara, Golfo Pérsico) están expuestas a neblina. Las regiones montañosas también son propensas a la presencia de neblina, debido a que las montañas representan un obstáculo en el movimiento de las masas de aire húmedo, obligándolas a elevarse por sus laderas y condensándose debido a enfriamiento adiabático. También, en regiones con clima templado, la neblina se puede formar debido a la condensación del aire húmedo caliente durante el día por el enfriamiento del suelo en la noche.

B. Dispositivos de colección de neblina

Existen dos alternativas para realizar la cosecha de neblina, puede ser tanto la intercepción de forma natural como de forma artificial. La intercepción de forma natural ocurre cuando el dosel de la vegetación (hojas y ramas que conforman su estructura) intercepta la neblina permitiendo el goteo del dosel hacia el suelo. En este caso, la colección se hace bajo la vegetación de forma directa. Este método se ha utilizado en sitios como en las Islas Canarias [55] y Omán [45], siendo ésta, la única fuente de agua tanto para sus habitantes como para animales [32]. Para este caso, la intercepción de la neblina va a depender de la estructura de la vegetación, de la forma de sus hojas y orientación respecto al viento. La cosecha de neblina implicaría la implementación de cisternas de recolección bajo la vegetación, lo cual puede ser complicado. También, las cisternas requerirían de un constante mantenimiento (para limpiar su superficie de ramas y hojas) y la infraestructura podría suponer una afectación al ecosistema (en el caso de bosques). La otra alternativa, es el uso de estructuras artificiales para captar neblina, como son los captadores de neblina. Los captadores tienen una malla que intercepta las pequeñas gotas de neblina a medida que el aire atraviesa la misma. Las gotas captadas en la malla se unen formando gotas más grandes que se escurren por la malla hacia una canaleta que conduce el agua a un tanque de almacenamiento.

Existen diferentes tipos de captadores de neblina, los cuales dependen de su estructura y tipo de malla. En cuanto a la estructura existen básicamente dos tipos: i) dispositivos cilíndricos (usado en Hawaii [56], Tenerife [52] y Costa Rica [57]), cuya malla constituye

la superficie de un cilindro, y el cual está conectado a un embudo que recoge la neblina captada, y ii) dispositivos bidimensionales (usado en Chile, España, Nepal, Sudáfrica y Australia [15], [39], [45]), que consiste en una superficie plana de malla ubicada perpendicularmente a la dirección del viento.

Los dispositivos cilíndricos tienen la ventaja de no ser dependientes de la dirección del viento, la desventaja es que no se pueden construir superficies muy grandes. Los dispositivos bidimensionales se pueden construir en varios tamaños dependiendo de las necesidades, no ocupan mucho espacio, son relativamente baratos y fáciles de instalar.

Sin embargo, tienen como desventaja que cuando existen vientos muy fuertes pueden ser desestabilizados o su malla puede romperse. Existen otro tipo de colectores menos usados, como: colector macrodiamante [58], atrapaniebla DISDERA [59], tipo Juvik [56], colector Eiffel [60] y two-stage tunnel tipo Daube [57]. Otra característica importante de los colectores de neblina es su tipo de malla. Tanto el material como su diseño influyen directamente en la eficiencia del colector. Un material común para las mallas es el acero, ya sea de aluminio [56], [57] o de acero inoxidable [15] con un diseño de pantalla de persiana o entrelazada con material de poliester. Este material ha sido probado efectivo durante eventos únicamente de neblina y eventos de lluvia y neblina con vientos fuertes [57].

Otro material utilizado tanto en colectores cilíndricos como en bidimensionales es el wire harp screen, que consiste en una red de hilos de nylon con un diámetro de 0.8 mm espaciados 2 mm [57], [61]. La desventaja de este material es su pérdida de eficiencia en eventos con fuertes vientos. El material mayormente utilizado es la red de sombra Raschel [15], que es de polipropileno y cuyo diámetro de fibra es efectivo para capturar la neblina. Además, el tejido consiste en triángulos estirados que beneficia el desplazamiento vertical por gravedad de las gotas que chocan contra ella [15], [45]. Uno de los colectores más usados alrededor del mundo (alrededor de 40 países [15]), es el "standard fog collector" (SFC) o colector de neblina estándar, propuesto por Schemenauer y Cereda 1994 [45]. Este colector consiste en un panel de red de sombra Raschel con

una superficie de recolección de $1m^2$, que está sostenido por una estructura en forma de marco a 2 m de elevación perpendicular a la dirección del viento.

La elección del tipo de colector de neblina va a depender de las condiciones del sitio, facilidad de construcción y su precio. Para el caso de la evaluación de la prefactibilidad de este sistema, se recomienda utilizar al menos dos colectores de diferentes materiales. Materiales conocidos y usados en otros sitios pueden servir para realizar comparaciones sobre la capacidad de colección de neblina del sitio de estudio y hacer predicciones previas sobre el potencial de captura a largo plazo en base a estas comparaciones. En el caso de que no se conozca las condiciones climáticas previas a la instalación de los colectores (en especial la velocidad y dirección del viento), se recomienda fuertemente el uso de al menos un colector cilíndrico. Aunque la cuantificación de la captura de neblina se podría realizar a nivel diario de forma manual, sería preferible realizar mediciones automáticas a un menor paso de tiempo para realizar comparaciones con otras variables meteorológicas.

C. *Influencia de factores climáticos*

Es importante identificar los factores climáticos que afectan la neblina. De acuerdo a la literatura, los principales factores que afectan la intercepción de neblina son el contenido de agua en el aire, el tamaño de las gotas y la velocidad del viento. Los dos primeros factores tienen relación con las propiedades de la neblina, mientras la velocidad del viento es un factor climático. La dirección del viento también puede ser un factor preponderante, la elección de usar un colector cilíndrico o bidimensional depende de esta variable. La lluvia es otro factor importante, sobre todo en regiones en donde existen precipitaciones del tipo orográficas, en donde las mismas están acompañadas de neblina. Además del viento y la lluvia, la temperatura, radiación solar y humedad relativa son otras variables meteorológicas que pueden afectar a la intercepción de neblina.

La implementación de una estación meteorológica en el sitio donde se planea evaluar la intercepción de neblina es fundamental. Se debe tratar que la estación mida al menos las principales variables

meteorológicas mientras se hace la cuantificación de la intercepción de neblina. El monitoreo de otras variables como la visibilidad y la distribución del tamaño de partículas de agua en suspensión sería un valor agregado, dado que se podría determinar el contenido de agua en el aire a partir de estos datos [62].

Para cuantificar el efecto de las variables meteorológicas en la intercepción de neblina se puede utilizar el coeficiente de correlación, *r(x,y)*. Aunque simple, el coeficiente de correlación además de reflejar la relación entre dos variables indica también el efecto multiplicador o reductor de la variable independiente sobre la dependiente. Otra forma de cuantificar este efecto es un análisis de cuantía, en donde se pueda determinar la importancia relativa de cada una de las variables climáticas en la recolección de neblina. Este análisis parte de la asunción de un modelo lineal múltiple, en donde existe una variable independiente que es función de variables independientes o predictores:

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n \text{ (1)},$$

donde y es la variable dependiente, $x1,2,..n$ son los predictores y $a1,2,\ldots,n$ son los parámetros de regresión. De acuerdo a la teoría de propagación de errores, el error de una variable en función de varias viene dado por la expresión:

$$\varepsilon(y) = \sum_{i=1}^{n} \left|\frac{\partial}{\partial x_i} y\right| \varepsilon(x_i) \text{ (2)},$$

donde n es el número de predictores, $\varepsilon(y)$ y $\varepsilon(x_i)$ son los errores de la variable dependiente y de los predictores, respectivamente. Para el caso de medidas directas (como las variables meteorológicas), el error de los predictores puede estar descrito por su desviación estándar (Sx_i). Usando la ecuación (1) como función en la ecuación (2), y resolviendo la ecuación (2), el error de la variable dependiente viene dado por:

$$\varepsilon(y) = a_1Sx_1 + a_2Sx_2 + \cdots + a_nSx_n \text{ (3)}.$$

Finalmente, si se normaliza el error de la variable independiente, se puede obtener la contribución relativa al error de cada una de las variables independientes (Px_i):

$$Px_i = \frac{a_i Sx_i}{\sum_{i=1}^{n} a_i Sx_i} \quad (4)$$

Para el caso del análisis de la intercepción de neblina, la variable dependiente corresponde a la intercepción de neblina, las variables independientes o predictores corresponden a las variables meteorológicas, y la contribución relativa al error de la variable dependiente de cada predictor se podría interpretar como el efecto relativo de cada variable meteorológica sobre la intercepción de neblina.

D. Separación de neblina y lluvia en el colector

Usualmente los diferentes dispositivos de colección de neblina atrapan tanto la neblina como la lluvia. En este caso no se tiene una medida del aporte real de la intercepción de neblina, una medida que es fundamental para la evaluación de la implementación de un sistema de recolección de neblina. En varios estudios [52], [56], [63], se utiliza una protección en la parte superior del colector para evitar la entrada de precipitación y medir únicamente la neblina. Sin embargo, los objetivos de estos estudios es la medición únicamente de la neblina y no el estudio de la factibilidad de un sistema de colección de neblina como es el caso de esta investigación.

La cuantificación de neblina es necesaria, por lo que una alternativa sería que además de los colectores de neblina se instale un dispositivo para realizar la medición únicamente de la neblina. Se debería considerar que el uso de un protector de lluvia en el dispositivo de colección de neblina, podría causar un efecto aerodinámico, alterando la dirección e intensidad del viento y por ende, afectando la estimación de neblina. Otra alternativa es realizar una separación de los aportes de neblina y lluvia en los colectores. Esta separación resulta ser una medida más directa del aporte de neblina ya que se la realiza directamente sobre el colector que se estaría evaluando y no indirectamente de otro dispositivo en donde se mide únicamente la neblina.

Para realizar la separación de neblina se puede utilizar un modelo geométrico sencillo. Domínguez *et al*. (2017) [43] utilizó este modelo para separar la neblina de la lluvia en un colector cilíndrico, al determinar primero el valor del volumen real de precipita-

ción captado por el colector de neblina, y calculando el volumen de neblina, como la diferencia entre el volumen medido por el colector y el volumen real de la lluvia (Figura 1).

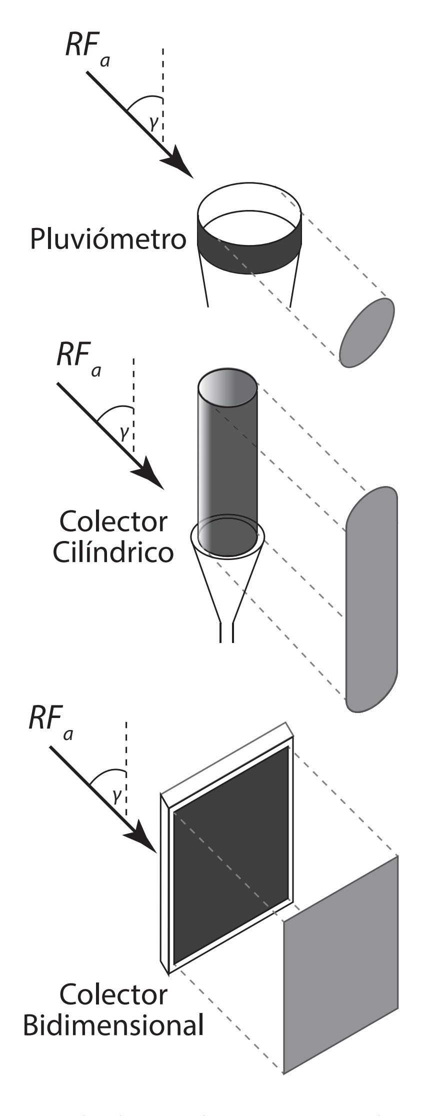


Figura 1. Área proyectada de un pluviómetro, colector cilíndrico y un colector bidimensional, perpendicular a la dirección de la precipitación

Esta metodología se basa en que la intensidad y dirección de la lluvia captada por un pluviómetro normal y el colector de neblina es la misma en ambos dispositivos. Esto puede ocurrir, si ambos dispositivos están ubicados cerca el uno del otro. También, es necesario asumir que la intercepción de neblina se produce sobre la proyección perpendicular a la dirección del viento de la malla del colector. En este caso, se necesitaría un medidor de velocidad y dirección del viento tridimensional. En el caso de tener un medidor estándar de la dirección del viento (únicamente mide la dirección horizontal del viento), la componente vertical de la dirección se podría asumir igual a la pendiente del terreno o de forma horizontal. Para casos de topografía sencilla y pendientes suaves se podría asumir una componente horizontal del viento. Pero en casos de topografía complicada o de pendientes fuertes se debería optar por la medición tridimensional o el uso de la pendiente del terreno, respectivamente. El modelo se describe a continuación.

Ante la existencia de viento (condición necesaria para que exista intercepción de neblina), el pluviómetro mide la componente vertical de la lluvia (RF_v) que viene dado de una relación trigonométrica:

$$RF_v = RF_a \times \cos\gamma \text{ (5)},$$

donde (RF_a) es la magnitud de la lluvia real que tiene una inclinación (angulo). El ángulo de inclinación de la lluvia se calcula a partir de la intensidad de la precipitación, el tamaño y velocidad terminal de caída de las gotas de lluvia, y de la velocidad del viento [40], [43]. La intensidad de la lluvia corresponde a la medición del pluviómetro directamente. El tamaño promedio del diámetro de una gota de lluvia (D) en mm puede ser calcula con [40], [43]:

$$RI = 2.23 \times (0.03937 \times RF_v)^{0.102} \text{ (6)},$$

donde *RI* es la intensidad de la lluvia en mm/h. La velocidad de caída de la gota (*U*) viene dado directamente de su diámetro [40], [43], [64]:

$$V = 3.378 \times \ln(RI) + 4.213 \text{ (7)},$$

y el ángulo de caída se calcula con [40], [43], [65]:

$$\tan\gamma = \frac{U}{V} \ (8),$$

donde *V* es la velocidad del viento.

A partir de la ecuación (5) se puede determinar la magnitud de la lluvia real, la cual es equivalente al volumen de lluvia real dividido para la proyección de la superficie circular horizontal del pluviómetro perpendicular a la dirección de RF_a. Realizando la misma equivalencia, el volumen captado por el colector de neblina correspondiente a la lluvia real (VR_a) es [43]:

$$VR_a = RF_a \times S_r \ (9),$$

donde S_r es la proyección de la superficie del colector perpendicular a la dirección de RF_a. Finalmente, la magnitud de la intercepción de neblina captada por el colector (*NI*) está dada por [43]:

$$NI = \frac{(VF_t - fe \times VR_a)}{S_v} \ (10),$$

donde S_v es la proyección de la superficie del colector perpendicular a la dirección del viento y es un coeficiente de corrección aplicado a la medición de la lluvia en el colector [43].

La dificultad de este método resulta en la obtención de las superficies proyectadas del colector de neblina, ya que va a depender de la dirección del viento y del ángulo de caída de la precipitación para cada paso de tiempo. En el caso de un colector bidimensional o cilíndrico, esta estimación no resulta difícil debido a que su geometría es relativamente sencilla, pero en el caso de geometrías complicadas (caso del colector tridimensional o colector con secciones complementarias) la estimación de esta superficie puede ser bastante desafiante.

3. Caso de estudio (San Cristóbal, Galápagos)

A. Área de estudio

1) Ubicación geográfica y clima

El Archipiélago de Galápagos está ubicado a 1000 km al oeste del Ecuador continental. La isla San Cristóbal se ubica al lado este del Archipiélago. La isla tiene una superficie de 558 km^2, donde más del 80% de su área está protegida por el Parque Nacional Galápagos.

La población total de la isla es de aproximadamente 7500 habitantes [66], donde la mayor concentración de la población está en Puerto Baquerizo Moreno. Las condiciones climáticas de Galápagos están influenciadas por la interacción de la corriente oceánica de Humboldt y los vientos alisios [67], [68]. Los vientos alisios, con dirección Sudeste, se condensan en las pendientes del barlovento de las islas principales generando una diferencia clara entre el lado húmedo barlovento y el lado seco sotavento de las islas. La migración interanual de la zona de convergencia intertropical da lugar a la aparición de dos temporadas climáticas, la fría de "garúa" y la cálida de "invierno". La temporada de garúa, ocurre entre junio y diciembre, y se caracteriza por la presencia de temperaturas relativamente bajas tanto del aire como de la superficie del océano [68]. Las precipitaciones son del tipo orográficas, de larga duración, baja intensidad y localizadas en el lado barlovento de las islas [68]. Durante esta época, una capa de inversión térmica se forma en la parte alta de las islas, dificultando el paso de las masas de aire húmedo y conduciendo de esta manera a la formación de neblina en las partes altas de las islas [69] La presencia de neblina es casi permanente durante esta época. Por otra parte, la temporada de invierno, ocurre de enero a mayo, y se caracteriza por la existencia de temperaturas relativamente altas en el aire y la superficie del océano. En esta época, ocurren precipitaciones del tipo convectivo, de corta duración y alta intensidad [68], [70], [71].

2) Geología y recursos hídricos

Las islas Galápagos, formadas de rocas basálticas mixtas [72]–[75], emergen de un punto caliente en una plataforma submarina poco profunda que forma la parte occidental de la cordillera Carnegie Ridge, en la placa de Nazca [75]. Pero la presencia de una cresta segmentada en el océano Centro de Propagación de Galápagos (GSC) no se puede ignorar, ya que afecta la estructura y organización de las islas. La isla San Cristóbal está formada por la erupción de flujos de lava pahoehoe y aa' [72] intercalado por periodos de calma, en donde debido a la condiciones húmedas en el clima, capas de suelo se desarrollaron debido a la meteorización [75]. Por este motivo, la estructura interna de la isla posee un sistema multicapas entre flujos de lava y paleo-suelos. Al final de la actividad volcánica, la isla fue cubierta

por una capa de depósitos de piroclastos de al menos 10 m [72], que debido a la meteorización se formó un suelo en la superficie.

En el lado barlovento de la isla, las condiciones húmedas y la presencia de suelo favorecieron la escorrentía promoviendo la erosión. Por este motivo se puede observar una red de drenaje desarrollado que tiene incisiones profundas en varios sectores. La hidrología es compleja en este lado de la isla, en la parte alta existen praderas con pendientes suaves que favorecen el encharcamiento de agua, como es el caso de la laguna "El Junco". A media elevación, se puede observar varias encañadas. Algunas secas y otras con agua permanente que nacen de fuentes que están siendo alimentadas por un acuífero colgado formado sobre una capa de baja permeabilidad que no permite la percolación del agua infiltrada a fondo [74]–[77]. El caudal de la mayoría de encañadas es menor a 5 l/s [78]. En la parte costera, el agua de las encañadas se vuelve a reinfiltrar, y solamente cuatro encañadas llegan a descargarse en el océano en la bahía de agua dulce [79].

Galápagos se caracterizan por la falta de agua en sus islas. La única isla con fuentes de agua dulce permanente es la isla San Cristóbal. El suministro de agua para la población es captado de dos encañadas, "La Toma" y "Cerro Gato", proporcionando alrededor de 30 l/s para la población de Puerto Baquerizo Moreno. Aunque actualmente las necesidades de agua para uso doméstico están satisfechas (aparentemente), existe un déficit importante en cuanto al riego. En el 2010, existió un déficit de 20.5% equivalente a 1.8 hm^3 [80], y se estima que este déficit se incremente a 7.1 hm^3 (80.1%) para el 2035 [81]. Este problema es aún mayor en otras islas poblados donde no existe fuentes de agua dulce permanente. Para garantizar las sostenibilidad de las poblaciones en estas islas, es necesario encontrar fuentes alternativas de agua, como es el caso de la cosecha de neblina.

B. Implementación de la metodología

1) Sitio de estudio

El sitio de estudio está ubicado en la parte alta del lado sur barlovento de la isla San Cristóbal, a 600 m.s.n.m (Figura 2).

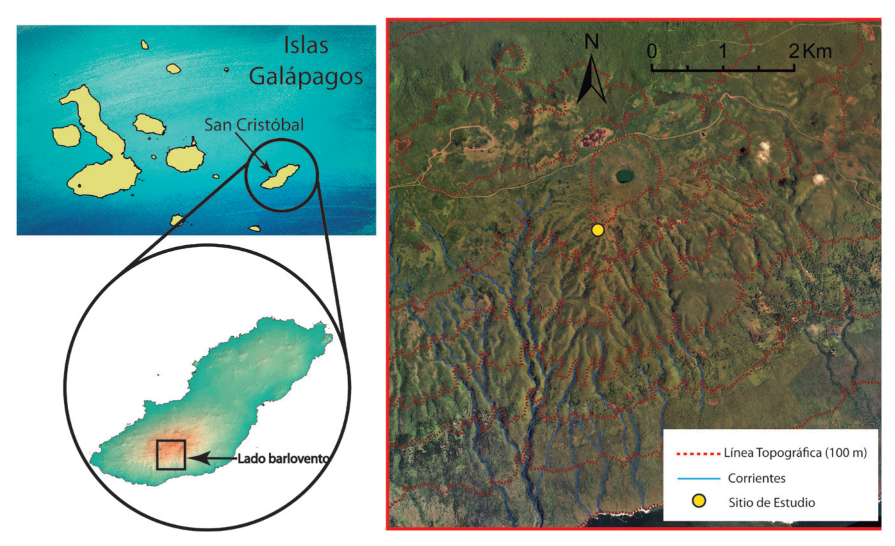


Figura 2. Ubicación del sitio de estudio en el lado barlovento de la Isla San Cristóbal (Galápagos)

Este sitio está dentro del área agrícola de la isla que se extiende desde la parte a media elevación de la isla (aproximadamente a 200 m.s.n.m) hasta la cumbre de la misma. En los alrededores del sitio, el suelo está cubierto por pasto y por bosques de *Miconia Robinsoniana*, una especie endémica de Galápagos. Toda esta zona (desde 300 m.s.n.m) está expuesta a la presencia de neblina alrededor de seis meses al año durante la época de garúa, por lo que representa un sitio ideal para evaluar y cuantificar la intercepción de neblina.

2) Monitoreo climático

Para realizar el monitoreo de las principales variables meteorológicas se utilizó una estación climática (Tabla 1).

Tabla 1. Equipos de la estación climática a 600 m.s.n.m. en San Cristóbal (Galápagos)

Equipo	Marca	Modelo
Datalogger[a] (x2)	Campbell	CR1000
Pluviómetro[b]	Texas	TE525MM
Pluviómetro[c] (x2)	Précis Mécanique	R307A0

Equipo	Marca	Modelo
Pluviómetro[d]	Campbell	ARG100
Sensor de temperatura y humedad relativa	Campbell	CS215
Anemómetro	Young	03002-5
Piranómetro	Campbell	CS300

a. Registrador de datos.
b. Usado para la medición de la precipitación.
c. Pluviómetro modificado para la medición de la intercepción de neblina en los colectores estándar (SFC35 y SFC50).
d. Pluviómetro modificado para la medición de la intercepción de neblina en el colector cilíndrico (CFC).

La precipitación se midió con un pluviómetro Texas TE525MM ubicado a 1.5 m del suelo. La dirección y velocidad del viento se monitoreó a 2 m sobre el nivel del suelo con un sensor Young WindSentry Kit. La radición solar se midió con un piranómetro de silicona Kipp&-Zonen SPLite a 2 m del suelo. La humedad relativa y temperatura se monitorearon con un sensor Campbell CS215 a 2 m sobre el suelo. Todos los datos de estas variables se grabaron en un datalogger Campbell CR1000, que permite realizar mediciones cada minuto y registra los datos cada 15 minutos.

La intercepción de neblina se midió utilizando tres colectores (Figura 3).

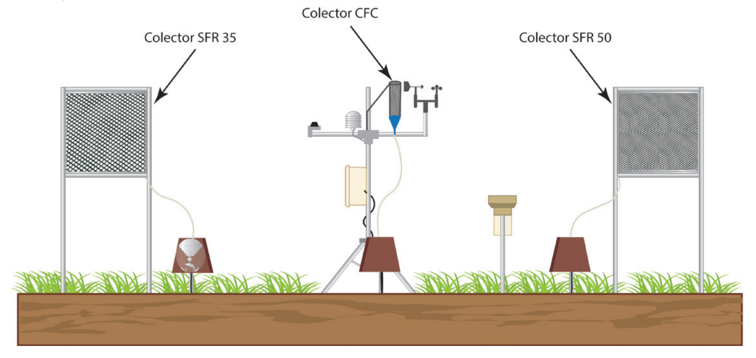


Figura 3. Esquema de monitoreo meteorológico y de intercepción de neblina. La estación meteorológica cuenta con un pluviómetro (1.5 m sobre el suelo), un anemómetro (2 m sobre el suelo), un sensor de temperatura y humedad relativa (2 m sobre el suelo), un piranómetro (2 m sobre el suelo) y tres colectores de neblina conectados a basculadores automáticos

El primero, es un colector cilíndrico (CFC), con una altura de 40 cm y 12 cm de diámetro, que tiene una estructura metálica cubierta por una malla plástica de 1 mm de apertura y un coeficiente de sombra del 72% (estimado a partir de análisis de imágenes). El agua interceptada por la malla se drena en un embudo en la parte inferior del colector, el cual está conectado a una manguera que dirige el agua hacia un sistema de basculación automático, el cual permite la medición del volumen de agua interceptado por el colector. Los otros dos dispositivos de medición de neblina, fueron los conocidos colectores estándar de neblina ubicados a 2 m sobre el nivel del suelo, con una superficie de colección de 1x1 m, y que tienen el mismo diseño pero difieren en la malla utilizada. Los colectores fueron construidos con malla de polietileno, de 35% de coeficiente de sombra en el primero (SFC35), mientras que el segundo colector tiene una malla de 50% de coeficiente de sombra (SFC50). La estructura de soporte de colector fue construida con tuberías de PVC, que resultó ser un material relativamente barato, de fácil transporte y que no afectó la sensibilidad del ecosistema de Galápagos. Estos colectores tienen una canaleta de recolección conectada a una manguera que dirige el agua captada por los colectores a sistemas de basculación automáticos individualmente. Los datos de los tres colectores fueron grabados por otro datalogger CR1000 que registra los datos de forma simultánea con los datos del resto de variables meteorológicas.

3) Aplicación de la metodología

La evaluación de la intercepción de neblina se realizó en los meses de junio y julio del año 2016, al inicio de la temporada de garúa. La cuantificación del rendimiento de los colectores para captar agua se lo realizó de dos maneras: i) con el volumen acumulado directamente medido por los colectores en l/m2·día, y ii) utilizando el modelo geométrico de separación entre intercepción de neblina y la lluvia a nivel horario, en mm. Para poder utilizar el modelo, se estimó el área proyectada de los colectores en el plano perpendicular a la dirección de la lluvia a nivel horario, en mm. Para poder utilizar el modelo, se estimó el área proyectada de los colectores en el plano perpendicular a la dirección de la lluvia (S_r) para cada paso de tiempo. En el caso del colector cilíndrico, S_r se

calculó en la forma descrita en Dominguez *et al*., 2017 [43], y para el caso de los SFCs, se utilizaron varias relaciones geométricas. Para todos los dispositivos, se asumió que el área de intercepción de proyección perpendicular a la dirección del viento es perpendicular a la pendiente del terreno. S_v fue estimado con las mismas relaciones trigonométricas utilizadas para obtener S_r, y es un valor constante para todo el periodo de estudio. El factor fe, correspondiente para la corrección de medición de la precipitación se estimó usando la metodología descrita en Domínguez *et al*., 2017 [43].

El análisis estadístico se realizó utilizando un filtro, primero los registros se discretizaron a datos horarios, usando la media o sumatoria de los datos cada 15 minutos. Luego, se utilizaron únicamente los datos en que existió un registro de intercepción de neblina diferente de 0, de esta manera se evalúa únicamente la relación entre las variables meteorológicas con la intercepción de neblina, cuando ésta ocurre.

4. Resultados

A. Condiciones climáticas durante el periodo de estudio

A partir de los datos históricos de la estación operada por la Fundación Charles Darwin en la costa de Santa Cruz, se puede observar que durante el periodo de estudio (junio y julio de 2016), la precipitación acumulada es de 26.5 mm y la temperatura media del aire es de 23.3 °C, que comparado, con los valores promedio de la estación (44.7 mm para la precipitación y 23.3 °C para la temperatura), resultan condiciones secas en términos de precipitación y condiciones normales en términos de temperatura.

Durante los meses de junio y julio de 2016 en el sitio de estudio la estación meteorológica registró las principales variables meteorológicas. Un resumen de esta información se presenta en la Tabla 2.

Tabla 2. Resumen de los valores diarios de las principales variables meteorológicas registras en el sitio de estudio

Variable	Valor
Precipitación[a] [mm]	72.0 (23.9)
Temperatura[b] [°C]	18.2 (15.5 a 24.0)

Variable	Valor
Radiación solar[c] [W/m^2]	211 (56 a 318)
Humedad relativa[d] [%]	99 (67)
Velocidad del viento[e] [m/s]	2.5 (1.9 a 3.1)
Dirección del viento[e] [^{O}]	199 (187 a 217)

a. Precipitación acumulada en todo el período corregida en función del ángulo de caída. En paréntesis, la precipitación máxima diaria registrada.
b. Temperatura media diaria. En paréntesis, el valor mínimo y máximo registrado.
c. Radiación solar media diaria durante las horas de sol. En paréntesis, el valor del primer y tercer cuartil.
d. Humedad relativa media diaria. En paréntesis, el porcentaje del tiempo que la humedad relativa fue mayor o igual a 99%.
e. Variable media diaria. En paréntesis, el valor del primer y tercer cuartil de la variable.

La precipitación registrada alcanzó un valor de 72 mm y la temperatura media fue de 18.2 °C. Estos valores resultan bajos comparando con los registros Junio y Julio del 2013 al 2016 en el mismo sitio (286.2 mm y 19.1 °C para la precipitación y temperatura, respectivamente). La velocidad y dirección media y diaria del viento fue de 2.5 m/s y 199°, respectivamente. La dirección del viento es similar a años pasados (193°), pero en el caso de la velocidad del viento, este valor está por debajo de los registrados en años pasados (2.9 m/s). La radiación solar media diaria durante las horas de sol permaneció relativamente baja durante todo el periodo (211 W/m^2) y es ligeramente mayor que la presentada en años pasados (202 W/m^2). Similar a años pasados, la humedad relativa estuvo cerca o igual a condiciones de saturación. La humedad relativa estuvo por encima del 95% en todo el periodo y durante el 67% del tiempo, ésta fue mayor al 99%.

B. Cuantificación de la intercepción de neblina y comparación entre colectores

Durante el periodo de estudio, el volumen total colectado por el colector de neblina CFC, SFC35 Y SFC50 fue de 26, 337 y 410 l, respectivamente. Estos valores representan un rendimiento diario de 3.8 l/m^2·día para el colector CFC, 7.0 l/m^2·día para el colector SFC35 y 8.5 l/m^2·día para el colector SFC50 (Tabla 3).

Tabla 3. Rendimiento e intensidad de intercepción de neblina de los tres colectores estudiados

Variable	CFC	SFC35	SFC50
Rendimiento[a] [$l/m^2 \cdot día$]	3.8 (0.6 a 4.9)	7.0 (2.0 a 10.0)	8.5 (2.2 a 12.1)
Intensidad[b] [mm/día]	3.7 (1.2 a 4.6)	4.1 (2.0 a 5.2)	4.8 (2.2 a 5.5)

a. Volumen de neblina y precipitación captada por el colector.
b. Altura de neblina captada por el colector como resultado del modelo geométrico de separación entre neblina y precipitación.
Los valores presentados entre paréntesis corresponden al primer y tercer cuartil de la variable.

El mayor volumen colectado a un nivel diario se registró en el colector SFC50, con un valor de 44.6 $l/m^2 \cdot$día. En todos los casos se logró colectar al menos 1 $l/m^2 \cdot$día durante el 77% del tiempo, y en el colector SFC50 durante el 88% del tiempo.

Se usó el modelo geométrico descrito anteriormente para separar la neblina de la precipitación. Los resultados del modelo muestran que la altura de intercepción de neblina durante el periodo de estudio fue de 176.5 mm, 195.8 mm y 230.6 mm en el colector CFC, SFC35 y SFC50, respectivamente. Estos valores representan una intensidad de intercepción de neblina media de 3.7 mm/día en el colector CFC, 4.1 mm/día en el colector SFC35 y 4.8 mm/día en el colector SFC50. La altura de intercepción de neblina presentó un alto porcentaje respecto a la precipitación, alcanzando el 245%, el 272% y el 320% de la precipitación en el colector CFC, SFC35 y SFC50, respectivamente.

Analizando los valores de intercepción de neblina a nivel diario, se realizó una correlación entre los volúmenes captados por los colectores, y también entre la altura de agua captado por los colectores (resultados del modelo). La correlación entre los tres colectores, tanto para el volumen como para la altura, es alta en todos los casos (Tabla 4).

Tabla 4. Correlación entre los colectores

Variable	CFC	SFC35	SFC50
CFC	1	0.87	0.95
SFC35	0.70	1	0.97
SFC50	0.76	0.96	1

Los valores con fondo blanco y gris representan el factor de correlación del volumen de neblina y precipitación captada entre colectores y la altura de neblina captada (resultado del modelo geométrico) entre colectores, respectivamente.

Los colectores estándar (SFC) presentan una mayor correlación entre sí que la que se presenta con el colector cilíndrico, sobre todo en el caso de la altura de neblina.

C. *Influencia de las variables meteorológicas en la intercepción de neblina*

Se realizó un análisis sobre la influencia de las variables meteorológicas en la intercepción de neblina de los colectores. En la Tabla 5 se muestra la correlación existente entre las variables meteorológicas y la intercepción de neblina en cada colector. La correlación es pobre para todos los casos. La variable que presenta mayor correlación con los tres colectores es la velocidad del viento (0.47 a 0.55). En el caso del colector CFC, la correlación con la precipitación también presenta un valor relativamente alto. En el caso de los colectores SCF, la otra variable que presenta una correlación relativamente alta es la radiación solar, aunque en este caso las correlaciones son negativas mostrando que esta variable tendría un efecto reductor en la intercepción de neblina. Para estos colectores (SFC), la correlación con la precipitación también es negativa, aunque la correlación es baja (-0.04 a -0.01).

Tabla 5. Correlación entre las variables meteorológicas y la intercepción de neblina en los colectores

Variable meteorológica	CFC	SFC35	SFC50
Precipitación	0.24	-0.04	-0.01
Velocidad del viento	0.45	0.47	0.55
Dirección del viento	0.09	0.09	0.09
Humedad relativa	0.05	0.11	0.04
Temperatura	0.07	0.05	0.16
Radiación solar[a]	0.05 (-0.03)	-0.18 (-0.29)	-0.07 (-0.19)

a. El valor entre paréntesis corresponde a la correlación entre la radiación solar y la intercepción de neblina únicamente durante las horas de sol

El análisis de cuantía se utilizó para determinar la influencia relativa de cada variable meteorológica en la variabilidad de la intercepción de neblina. Los resultados se muestran en la Tabla 6.

Tabla 6. Contribución relativa de las variables meteorológicas en la variabilidad de la intercepción de neblina en los colectores

Variable meteorológica	CFC	SFC35	SFC50
Precipitación	22.6	8.5	11.6
Velocidad del viento	54.2	49.1	54.5
Dirección del viento	5.2	1.1	2.8
Humedad relativa	5.0	3.1	1.1
Temperatura	2.3	4.8	7.1
Radiación solar	10.7	33.4	22.9

Los valores de la contribución de cada variable en la intercepción de neblina se presentan como porcentaje.

Las variables que mayor influencian tienen en la intercepción de neblina son la velocidad del viento, radiación solar y la precipitación. La velocidad del viento es la variable que mayor influencia tiene en la intercepción de neblina de los tres colectores, con una contribución aproximada del 50%. En los colectores estándar, se puede apreciar que la radiación solar tiene una contribución importante, sobre todo en el colector SFC35, alcanzando un valor del 33%. En estos colectores, la precipitación podría ser considerados como un factor influyente en menor medida, alcanzando un valor del aproximado del 10%. A diferencia de los colectores estándar, en el colector cilíndrico la intercepción de neblina está mayormente influenciado por la precipitación (22.6%) que por la radiación solar (10.7%).

5. Discusión

A. *Influencia del clima*

El análisis de cuantía determinó la influencia de las principales variables meteorológicas. La velocidad del viento se consolidó como la variable que mayor influencia tiene sobre la intercepción de neblina, alcanzando el 54.2%, 49.1% y el 54. 5% en el colector CFC, SFC35 y SFC50, respectivamente. Estos resultados coinciden con otros estudios en donde se asevera que la velocidad del viento es la variable más influyente en la recolección de agua de neblina [54], [82] (Ritter *et al*., 2008; Schemenauer & Joe, 1989). Aunque la correlación directa entre la velocidad del viento y la intercepción

de neblina en los colectores no es buena (< 0.55), la influencia puede ser evidenciada con un gráfico de caja-bigote (Figura 4).

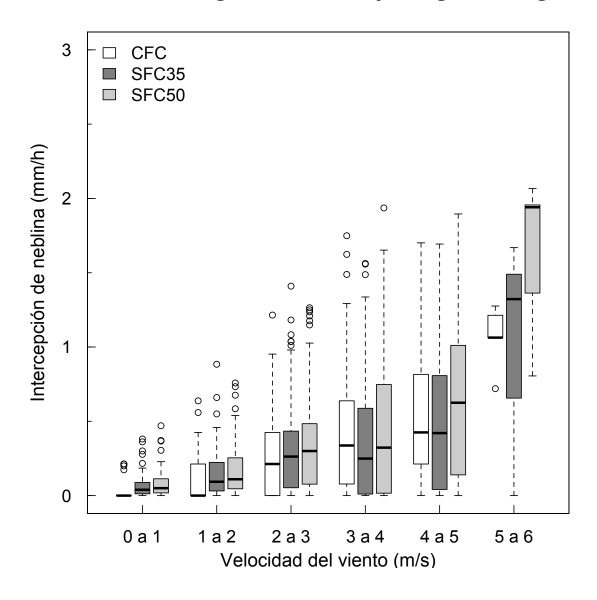


Figura 4. Estadísticas de la intensidad de intercepción de neblina en cada colector clasificadas por intervalos de velocidad del viento durante el periodo de estudio (Junio-Julio 2016)

En los tres colectores, es evidente que la intercepción de neblina incrementa conforme la velocidad del viento incrementa, sobre todo en los casos de velocidades del viento altas (mayores a 3 m/s). En todos los rangos de velocidad, se observa que existió intercepción de neblina en los tres colectores. La intercepción de neblina puede ocurrir por dos procesos principales, deposición e impactación [83]. En condiciones de velocidad del viento baja, la intercepción de neblina es controlada por deposición principalmente, por lo que es necesario que el contenido de agua en la neblina sea relativamente alto y que la superficie de intercepción sea eficiente. Con velocidad del viento alta, la impactación es el principal mecanismo de intercepción de neblina dado que el contenido de agua en la neblina es inversamente proporcional a la velocidad del viento debido al incremento de la evaporación potencial [83]. Esto quiere decir que superado cierto

valor de velocidad del viento el contenido de agua en la neblina se va a reducir al punto en que la intercepción de neblina se reduzca. De hecho, Villegas *et al.*, 2008 [83] concluye que las condiciones más óptimas de intercepción de neblina son cuando existe un contenido de agua en la neblina promedio y velocidades del viento promedio. En este caso de estudio, se puede apreciar que el contenido de agua en la neblina es relativamente alto, dado que no se alcanzó el umbral de velocidad del viento en donde la intercepción de neblina se reduzca. Se podría deducir que este hecho evidencia el potencial de intercepción de neblina en la parte alta de San Cristóbal, donde incluso con una reducción en el contenido de agua en la neblina (efecto que podría presentarse debido al cambio climático) todavía se tendría condiciones óptimas de intercepción de neblina.

La dirección del viento no tuvo una influencia importante en la intercepción de neblina de los colectores, incluyendo los bidimensionales (SFC35 y SFC50). Esto puede ser explicado porque la dirección del viento se mantuvo relativamente constante durante todo el periodo (Figura 5).

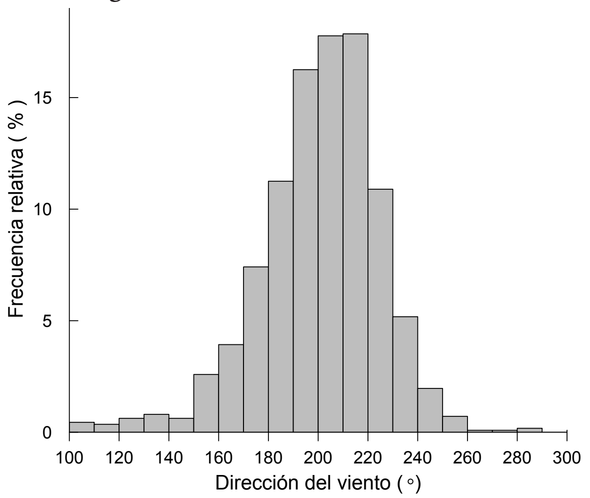


Figura 5. Diagrama de frecuencias relativas de la dirección del viento en todo el período de estudio, Junio-Julio 2016

Durante el 90% del tiempo en que se registró intercepción de neblina en alguno de los colectores, la dirección del viento se mantuvo entre 170 y 230°, que correspondería apenas en una reducción máxima del 13% de la intercepción de neblina en la dirección perpendicular al área de colección de los colectores. Estas condiciones son ideales para el uso de un colector bidimensional sobre un cilíndrico, siendo el primero más sencillo de construir, ocupa menos espacio para una misma área de colección y resulta más barato (en la siguiente sección se discutirá su eficiencia).

La radiación solar también presentó una influencia importante en la intercepción de la neblina en todos los colectores. Su contribución en la intercepción de neblina es la más variable alcanzando un valor de 10.7% en el colector CFC, 33.4% en el colector SFC35 y 22.9% en el colector SCF50. Existe una influencia apreciable entre la magnitud de la radiación solar durante las horas de sol y la intensidad de la intercepción de neblina por los colectores (Figura 6), siendo la intercepción de neblina por los colectores mayor cuando la radiación solar es más baja. Por esto, la radiación solar puede ser una variable que evidencie la presencia de neblina y de su contenido de agua en el aire.

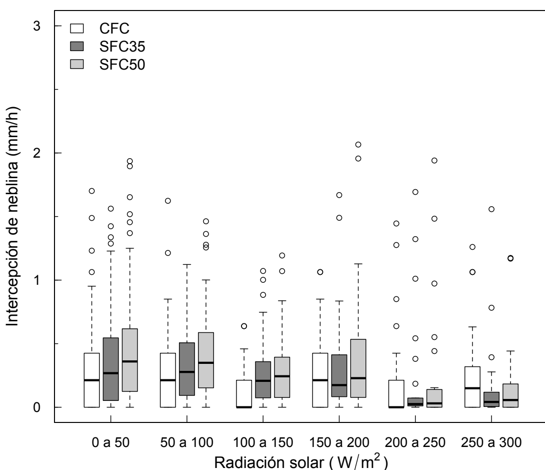


Figura 6. Estadísticas de la intensidad de intercepción de neblina en cada colector clasificadas por intervalos de radiación solar durante el periodo de estudio (junio-julio 2016)

La intercepción de neblina durante las horas del día ocurrió durante el 38%, el 39% y el 41% del tiempo en el colector CFC, SFC35 y SFC50, respectivamente. Esto quiere decir que durante mayor tiempo (aproximadamente el 60%), la intercepción de neblina ocurrió en la noche. La radiación solar promedio cuando existió intercepción de neblina en alguno de los colectores fue de 133 W/m^2. Este valor resulta bajo comparado con la radiación solar promedio cuando no existió intercepción de neblina (285 W/m^2) y peor con la radiación solar promedio para cielo despejado (441 W/m^2) calculado a partir de la radiación extraterrestre. Sin embargo, esto puede ser debido a que la intercepción de neblina pudo ocurrir en horarios muy tempranos o muy tarde del día. Esto no es correcto ya que la intercepción de neblina ocurrió durante todo el día (Figura 7), presentándose las mayores magnitudes de intercepción de neblina en los colectores alrededor de las 7 de la mañana y las menores magnitudes alrededor de las 1 de la tarde. La reducción de la radiación solar debido a la presencia de neblina ya se ha evidenciado en otros sitios, como es el caso del lado sotavento de Hawaii, en el cual la neblina redujo en 36% la insolación, en Puerto Rico el 48% y en las Islas Canarias el 58% [54], [84], [85]. En el caso de San Cristóbal, la reducción de la insolación fue relativamente baja alcanzando un valor del 30%.

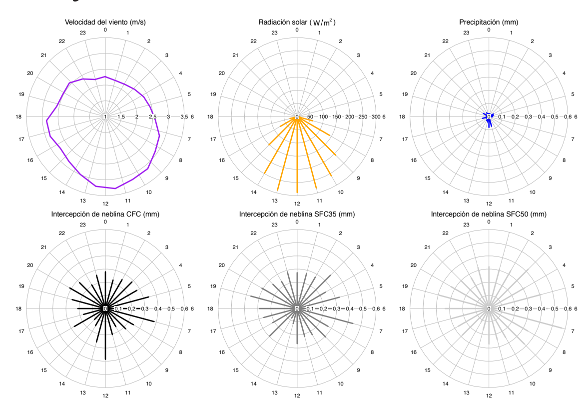


Figura 7. Variación horaria de las variables meteorológicas más influyentes en la intercepción de neblina (velocidad del viento, radiación solar y precipitación) e intercepción de neblina por cada colector durante el periodo de estudio (Junio-Julio 2016)

La precipitación es otra de las variables que afecta la intercepción de neblina pero en menor medida. Su contribución en la intercepción de neblina fue del 22.6%, 8.5% y 11.6% en el colector CFC, SFC35 y SFC50, respectivamente. Del total de neblina interceptada por los colectores, la neblina captada sin la existencia de precipitación representó el 75.7% en el colector CFC, el 98.2% en el colector SFC35 y el 95.7% en el colector SFC50. Claramente la precipitación no tiene influencia en la intercepción de neblina de los colectores estándar, pero sí la tiene en el colector cilíndrico. Al ver el gráfico de caja-bigote (Figura 8), se puede evidenciar que la intensidad de intercepción de neblina del colector CFC no está influencia por la intensidad de precipitación, y que en este colector se llega interceptar neblina con las mayores intensidades de precipitación (0.3 y 0.4 mm/h). Aunque el porcentaje de neblina interceptaba con intensidades mayores a 0.2 mm/h es de apenas el 12%.

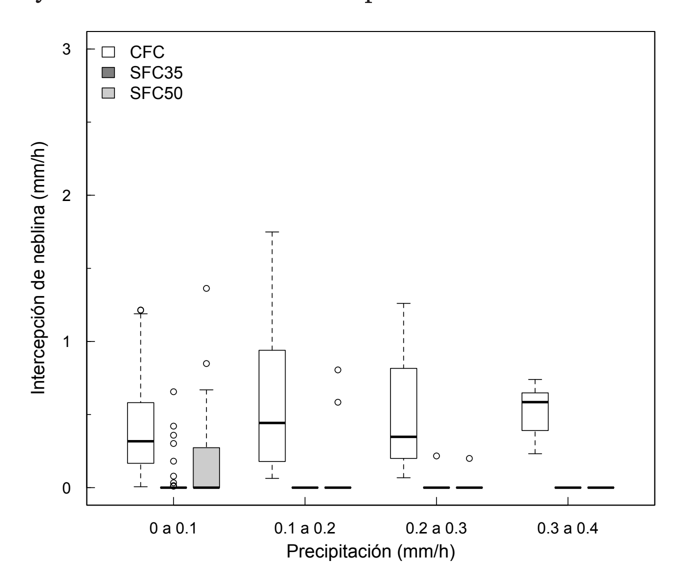


Figura 8. Estadísticas de la intensidad de intercepción de neblina en cada colector clasificadas por intervalos de precipitación durante el periodo de estudio (Junio-Julio 2016)

Variables como la temperatura y la humedad relativa presentan una contribución pequeña en la intercepción de neblina (con-

tribución combinada menor del 8.2% en el mejor de los casos que corresponde al colector SFC50). Sin embargo, esto no quiere decir que su monitoreo no sea importante o que las variables no tengan realmente influencia. Ambas variables están relacionadas con el contenido de agua en el aire de la neblina. La humedad relativa indica el grado de saturación del aire, por lo que mientras mayor es la humedad relativa, mayor es el contenido de agua en el aire. La temperatura también tiene influencia en esta variable ya que el contenido de agua en el aire de la neblina está inversamente relacionado con el calor latente de condensación de la temperatura del agua [86]. Dominguez *et al.*, 2017 [43] al estudiar la intercepción de neblina por parte de la vegetación endémica de San Cristóbal, compararon la intercepción de neblina por parte de la vegetación de dos temporadas de garúa distintas, y evidenciaron una diferencia significante, que dadas las condiciones (existió una diferencia significativa en la temperatura de las dos temporadas producto de una anomalía en la temperatura de la superficie del mar por el efecto temprano del evento de El Niño en la región 1+2), la atribuyeron a una diferencia en el contenido de agua en el aire de la neblina.

En este caso de estudio, el monitoreo de la temperatura y la humedad relativa fue de dos meses durante la temporada de garúa, en donde las condiciones fueron relativamente constantes y es por esto que las dos variables no tuvieron una gran influencia en la intercepción de neblina por parte de los colectores. Posiblemente con un monitoreo más largo, que incluya la temporada cálida de invierno, se podría ver una mayor influencia de estas variables.

B. Comparación de colectores

El colector que presenta la mayor eficiencia de captación de neblina es el SFC50. Su intensidad de intercepción de neblina durante el periodo de estudio (4.8 mm/día) es 1.31 veces mayor que el colector CFC y 1.18 veces mayor que el colector SFC50. En el caso de captación de precipitación y neblina, también presenta un mayor rendimiento. El colector CFC presenta la peor eficiencia para captación de neblina. Esto puede ser explicado al repasar la dinámica del flujo de las gotas de agua de neblina que atraviesan

una malla. La eficiencia de colección de neblina de un colector depende de su eficiencia aerodinámica, eficiencia de impactación y eficiencia de drenaje [61], [87]–[89]. La eficiencia de aerodinámica está relacionada a la capacidad que tiene la malla de dejar pasar el flujo de aire con neblina a través del mismo sin perturbar su curso. La eficiencia de impactación depende de la capacidad que tiene la superficie de la malla en atrapar las gotas de agua en la neblina que atraviesa la misma. La eficiencia de drenaje se refiere la capacidad que tiene la malla en drenar el agua de neblina captada por la malla rápidamente. La eficiencia total de un colector va a depender de la interacción de estas tres eficiencias. El colector CFC al ser cilíndrico, tiene una superficie aerodinámica que ayuda al paso lateral del flujo de aire con neblina, por lo que una parte del flujo en vez de atravesar el colector pasa rodeando su superficie y por lo tanto reduciendo su capacidad de colectar neblina. Además, su coeficiente de sombra relativamente alto, su pequeño tamaño de las aperturas de la malla y la forma cruzada de sus elementos, a pesar de mejorar la eficiencia de impactación (al existir una mayor área de obstrucción) producen un efecto adverso en la intercepción de neblina, las gotitas de agua que se interceptan bloquean el flujo de aire a través de los agujeros de la malla volviéndola más impenetrable [61]concomitantly with a reduction in ambient temperature and an increase in wind velocity (u, produciendo que una mayor parte del flujo de aire no atraviese la malla y se desvíe. Este flujo que es desviado además de no llevar las gotas de agua a la malla puede acarrear las gotas que se encuentran en la malla y sacarlas fuera del área de colección. Esta situación podría empeorarse debido a la forma de la malla, que al tener fibras horizontales a poco distancia, la eficiencia de drenaje es reducida [88]. Este conjunto de factores podría explicar el bajo rendimiento del colector CFC en comparación de los otros colectores y además el por qué la intercepción de neblina en este colector es influenciado por la precipitación. En los eventos de neblina y precipitación, probablemente la lluvia, al tener un diámetro de gotas más grandes, choca con la malla y coalesce con las gotitas de agua de neblina interceptadas por la malla, mejorando el drenaje como consecuencia de su mayor tamaño.

Los dos colectores SFC tienen las mismas dimensiones y la misma estructura. La única diferencia es su coeficiente de sombra. El SFC35, al tener más aberturas que el SFC50 permite el flujo del aire con neblina a través de su malla más fácilmente reduciendo la resistencia aerodinámica. Por el contrario, SFC35 tiene una menor área de impactación reduciendo la posibilidad de captar gotas de neblina en cuyo caso SFC50 sería más eficiente. Determinar la eficiencia de impactación y drenaje de estos colectores resulta en una serie de ecuaciones que dependen de varias variables y cuya determinación no entran en el alcance de este estudio (se recomienda leer Regalado & Ritter, 2016 [88] para revisar estas ecuaciones). Sin embargo, la eficiencia aerodinámica puede ser determinada de forma teórica. Rivera, 2011 [87] utilizó un principio de superposición para separar el flujo de aire que elude un colector de neblina teórico impermeable y el flujo que pasa libremente a través de los agujeros de un colector de neblina. A partir de estas relaciones obtuvo el siguiente modelo para determinar la eficiencia aerodinámica:

$$\eta_A = \frac{SC}{1+\sqrt{C_O/C_D}} \quad (11),$$

donde *SC* es el coeficiente de sombra de la malla, C_D es el coeficiente de arrastre y C_O es el coeficiente de caída de presión. El coeficiente de arrastre depende del número de Reynolds y para flujo turbulento (esperado en condiciones típicas de neblina), *C* tiene valores desde 1 para colectores cilíndricos hasta 2 para colectores cóncavos [87], [88]. Colectores de forma cuadrada alcanzan el valor de 1.18 [87], [90]. Co resulta de la relación de presiones entre el lado barlovento y sotavento del colector. Para el caso de una malla de seda (lo más parecido al material plástico o polietileno), puede ser determinado por [87]:

$$C_O = 1.62\left(1.3SC + \left(\frac{SC}{1-SC}\right)^2\right) \quad (12),$$

siendo C_O únicamente dependiente del coeficiente de sombra. De acuerdo a estas fórmulas, la eficiencia aerodinámica del colector CFC, SFC35 y SCF50 es de 0.16, 0.17 y 0.2, respectivamente. La eficiencia aerodinámica del colector SFC50 es la mayor, coincidiendo con los resultados encontrados en este estudio. La relación de las

eficiencias aerodinámicas entre colectores es similar a la relación del rendimiento de captación de neblina entre colectores (Tabla 7).

Tabla 7. Relación de rendimiento experimental y eficiencia aerodinámica teórica entre colecotres

Relación	Rendimiento experimental[a]	Eficiencia teórica[b]
SCF50/SFC35	1.18	1.15
SCF50/CFC	1.31	1.25
SFC35/CFC	1.11	1.09

a. Rendimiento de cada colector estimado a partir de la intercepción de neblina total registrada en el periodo de estudio.
b. Eficiencia teórica de cada colectore estimada a partir de la ecuación 11.

De acuerdo a la literatura [82], [87], [88] el colector de 50% de coeficiente de sombra tiene una mayor eficiencia que el colector de 35%. De acuerdo al modelo usado (Figura 9), las mejores eficiencias aerodinámicas se presentan entre valores de 50 y 60% de coeficiente de sombra [87]. Estos valores coinciden con las conclusiones de Schemenauer & Joe, 1989 [82] en un estudio experimental en el norte de Chile, en donde usó una doble capa de malla Raschel de 35% de coeficiente de sombra (que resulta en valores aproximados al 50% de coeficiente de sombra).

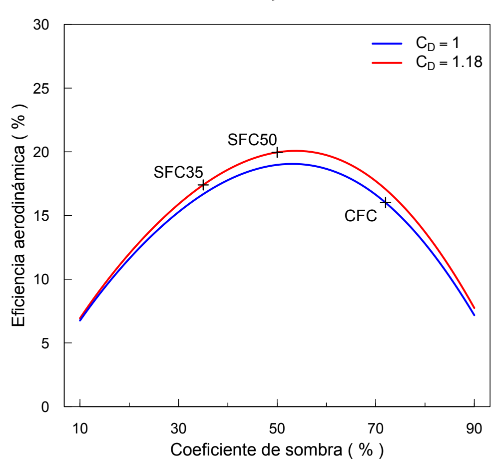


Figura 9. Variación de la eficiencia aerodinámica en función del coeficiente de sombra de un colector (ecuación 11)

C. Comparación con otros sitos

En San Cristóbal el colector SFC50 tuvo un rendimiento de captura de neblina y precipitación de 8.5 l/m^2·día durante dos meses de la época de garúa. Este valor se encuentra por encima del promedio mundial (Figura 10).

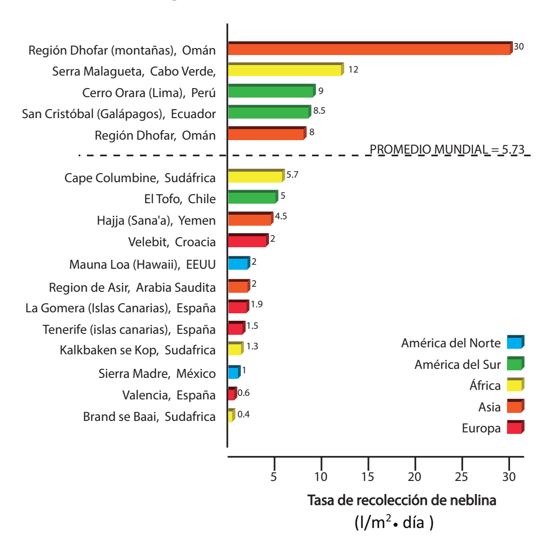


Figura 10. Tasa de intercepción de neblina en varios sitios a nivel global

Se han reportado rendimientos de 3 y 9 l/m^2·día, en Chile y Perú, respectivamente, durante la época de neblina utilizando colectores de neblina estándar [45]. En un estudio en la zona del Caribe sudamericano (Colombia y Venezuela), Cavelier & Goldstein, 1989 [37] realizaron un estudio sobre la intercepción de neblina encontrando que la intercepción de neblina varía entre 0.2 y 2.2 mm/día e identificaron que la intercepción de neblina sigue un gradiente orográfico en esta zona. En Sierra Madre, México, se registran valores entre 0.4 y 1.6 l/m^2·día a elevaciones de 1330 a 1900 msnm [91]. Mientras que la recolección en Mauna Loa en Hawaii, a 2530 y 1580 msnm, alcanza valores de 1.9 y 2.1 mm/día utilizando un colector cilíndrico tipo Juvik [56]. En Asia, en regiones áridas, se han realizado mediciones de intercepción de neblina con varios propósitos. En Omán, Yemen y Arabia Saudita la recolección

de neblina ha sido de 8, 4.5 y 2 l/m^2·día, respectivamente [45], [82], [92]. En las montañas de la región de Dhofar (Omán), en una elevación entre 900 y 1000 msnm, se registró una alta recolección de neblina con un promedio de 30 l/m^2·día durante tres meses de la temporada del monzón [45]. En el caso de Europa, se han realizado estudios principalmente en las islas Canarias, encontrando valores entre 0.2 y 5.0 l/m^2·día en la isla La Gomera [54] y valores de 1.3 a 1.8 l/m^2·día en Tenerife [52], [61]Taborno (1015 m a.s.l.. En la región Mediterránea de Valencia, se registró valores máximos de recolección de neblina hasta 7 l/m^2·día [51]mientras que en Croacia en la montaña Velebit, la intercepción de neblina alcanzó valores hasta de 4 l/m^2·día [93]. En África, la recolección de neblina también ha sido relevante. Sitios como la costa oeste de Sudáfrica, donde se han recolectado 0.4 l/m^2·día y 1.3 l/m^2·día, en Brand se Baai y Kalkbaken se Kop, respectivamente [94]. En éste último, la recolección de neblina es mayor y se atribuye a que se encuentra a más altura y además la velocidad del viento es mayor, siendo de 5.2 m/s en promedio, mientras que en Brand se Baai es de 4.3 m/s. En Cape Columbine el promedio de recolección de agua es de 5.7 l/m^2·día [94]. En Cabo Verde, se reportó un promedio de recolección de neblina de 12 5 l/m^2·día [95].

D. Implementación del sistema de colección de neblina

De acuerdo a los resultados de este estudio, se pudo evidenciar que tanto los colectores estándar como el colector cilíndrico son propicios para la captura de neblina en el Archipiélago. Dado que la dirección del viento no sufre variaciones drásticas y que los colectores estándar tienen un mejor rendimiento, se recomendaría el uso de colectores estándar, de preferencia colectores SFC50. Echeverría, 2018 [96] estimó la demanda de agua para consumo doméstico de la zona rural de San Cristóbal, obteniendo un valor de 80.3 l/día·habitante. Para el caso de fincas, la autora estimó que una familia promedio se compone de 5 personas, por lo que una finca tendría una demanda de 401.5 l/día. Esto quiere decir que de acuerdo a los resultados de este estudio, durante la época seca, colectores de neblina con una malla raschell de 50% de coeficiente de sombra con

un área de intercepción de 48 m^2 ubicados a una elevación de 600 m.s.n.m. podrían cumplir 100% de la demanda (para elevaciones más bajas se necesitaría un área mayor de recolección). Utilizando un sistema de colección de aguas lluvias en el techo con un área de 200 m2 se podría colectar 0.3 l/día. Existen varias comunidades que han utilizado la cosecha de neblina como una fuente de agua para su población. Chungungo, una comunidad ubicada en la zona desértica de la costa de Chile utiliza un sistema de cosecha de neblina de 100 colectores grandes (10x4 m), produciendo en promedio 15000 l de agua potable para 300 habitantes de la villa [32]. De acuerdo a Batisha, 2015 [15], otros proyectos de colección de neblina están produciendo 6300 l/día en Tojquia (Guatemala) con 1400 m2 de superficie de colección y cantidades viables en sitios como Sudáfrica, Perú y Colombia.

Los sistemas de colección de neblina además de ser amigables con el medio ambiente (no requieren energía eléctrica), sostenibles en el tiempo (bajo costo operacional) también resultan ser relativamente baratos comparados con otros sistemas convencionales que requiere una inversión inicial grande, costos operacionales altos y un alto mantenimiento [32].

En Chile, un sistema de cosecha de neblina puede ser producido por cerca de $1 por m^3 y transportado a los hogares de una comunidad lejana por $1 por m^3 [44]. Batisha, 2015 comenta que el costo de 100 colectores de neblina grandes, adecuado para una pequeña comunidad, estaría en el orden de $40000, que es barato comparado con otros sistemas de agua. Claramente, este sistema alternativo de agua tiene un potencial alto. Ciudades y comunidades inteligentes, deben tomar en cuenta sistemas de agua sustentables y amigables con el ambiente como es el caso de la cosecha de neblina.

6. Conclusiones

La parte alta de la isla San Cristóbal (Galápagos) está cubierta por una capa de neblina casi permanente durante 7 meses al año. Aprovechando este fenómeno, durante dos meses se ha instalado colectores de neblina para cuantificar la intercepción de neblina y evaluar su potencial como una fuente de agua alternativa para las

comunidades del sector. Se usaron tres colectores de neblina, uno cilíndrico y dos colectores estándar bidimensionales de 35 y 50% de coeficiente de sombra. Las mediciones fueron complementadas con el monitoreo de las principales variables meteorológicas.

Los resultados muestran que el rendimiento de los colectores es de 3.8, 7.0 y 8.5 l/m^2·día en el colector CFC, SFC35 y SCF50, respectivamente. Utilizando un modelo geométrico, se determinó la profundidad de intercepción de neblina únicamente (separado de la precipitación), obteniéndose valores de 176.5 mm en el colector CFC, 195.8 mm en el colector SFC35 y 230.6 en el colector SFC50. Estos valores corresponden a más del 242% de la lluvia, lo que evidencia el potencial de intercepción de neblina en el sector. Así mismo, el rendimiento del colector SFC50 está por encima del promedio de intercepción de neblina mundial.

Al igual que en otros estudios, se determinó que la velocidad del viento es la variable que más influencia tiene sobre la intercepción de neblina (aprox. 50%), seguida de la radiación solar y de la lluvia. Esta última tiene una mayor influencia en el colector CFC. El hecho de que la intercepción de neblina incremente con la velocidad del viento y no se estanque en una meseta indica el potencial de neblina de sector, incluso con velocidades del viento más fuertes o un contenido de agua en el aire más bajo. La dirección del viento no tuvo influencia sobre la intercepción de neblina porque permaneció relativamente constante durante el periodo de estudio, lo que justifica el uso de colectores bidimensionales en vez de colectores cilíndricos que tienen una menor eficiencia de recolección de neblina.

El colector SFC50 resultó ser el que mayor eficiencia de intercepción de neblina tiene. Esto se puede explicar por su coeficiente de sombra, el cual permite una mayor área de obstrucción para interceptar las gotas de agua en el flujo de aire que atraviesa su malla sin causar demasiadas obstrucciones que obliguen al flujo de aire desviarse de su trayectoria de intercepción con la malla. El colector CFC es el que menor eficiencia tiene, debido a su forma aerodinámica permite el paso lateral del flujo del aire. Además, el pequeño tamaño de sus aberturas induce a la obstrucción de las mismas con gotas de agua, haciendo que la superficie sea menos permeable y por lo tanto se capte menor neblina. Se concluye que la influencia de la lluvia en

la intercepción de neblina en este colector es debido a que ayuda al drenaje rápido de la neblina interceptada en la malla.

El potencial de cosecha de neblina en Galápagos como una fuente alternativa de agua resulta evidente. Este sistema podría ser implementado como una fuente alternativa de agua para este sector, sobre todo en las islas donde los recursos hídricos superficiales permanentes son escasos o inexistentes. Esta información es de vital importancia para las autoridades que trabajan en el manejo de los recursos hídricos en estas islas, sobre todo por el hecho de la escasez de fuentes superficiales de agua en este territorio. Así mismo, la metodología usada sirve para la evaluación de la implementación de un sistema de cosecha de neblina en otras partes del mundo. Claro que sería necesario complementar esta investigación con la evaluación de la calidad del agua interceptada de la neblina, una evaluación económica y el uso de otras tecnologías de intercepción de neblina (ej. material hidrofóbico [97]). En la planeación de desarrollar ciudades inteligentes, la sostenibilidad es un aspecto importante. Técnicas como la cosecha de neblina, entran dentro de esta idea de sostenibilidad.

Agradecimientos

El estudio se llevó a cabo dentro del marco del proyecto Galápagos Islands Integrated Water Studies (GIIWS). El equipo GIIWS quisiera agradecer a la Agence Nationale de la Recherche (ANR-blanc 2010 GIIWS Ref.601-01) y Galapagos Conservancy por su financiamiento. También quisiéramos agradecer a nuestros colaboradores locales en Galápagos: la Estación de Investigación Charles Darwin, el Parque Nacional Galápagos, y propietarios privados que permitieron la instalación de las estaciones climáticas.

Referencias

[1] V. Albino, U. Berardi, and R. M. Dangelico, “Smart Cities: Definitions, Dimensions, Performance, and Initiatives,” *J. Urban Technol.*, vol. 22, no. 1, pp. 3–21, Jan. 2015.

[2] T. Bakıcı, E. Almirall, and J. Wareham, “A Smart City Initiative: the Case of Barcelona,” *J. Knowl. Econ.*, vol. 4, no. 2, pp. 135–148, 2013.

[3] T. M. Chen, "Smart grids, smart cities need better networks [Editor's Note]," *IEEE Netw.*, vol. 24, no. 2, pp. 2–3, 2010.
[4] L.-G. Cretu, "Smart Cities Design using Event-driven Paradigm and Semantic Web," *Inform. Econ.*, vol. 16, no. 4, pp. 57–67, 2012.
[5] C. Harrison *et al.*, "Foundations for Smarter Cities," *IBM J. Res. Dev.*, vol. 54, no. 4, pp. 1–16, 2010.
[6] P. Lombardi, S. Giordano, H. Farouh, and W. Yousef, "Modelling the smart city performance," *Innov. Eur. J. Soc. Sci. Res.*, vol. 25, no. 2, pp. 137–149, Jun. 2012.
[7] G. C. Lazaroiu and M. Roscia, "Definition methodology for the smart cities model," *Energy*, vol. 47, no. 1, pp. 326–332, 2012.
[8] L. Guan, "Smart Steps To A Battery City," *Goverment news*, vol. 32, no. 2, pp. 24–27, 2012.
[9] N. Komninos, "Intelligent cities: Variable geometries of spatial intelligence," *Intell. Build. Int.*, vol. 3, no. 3, pp. 172–188, Jul. 2011.
[10] K. Kourtit, P. Nijkamp, and D. Arribas, "Smart cities in perspective – a comparative European study by means of self-organizing maps," *Innov. Eur. J. Soc. Sci. Res.*, vol. 25, no. 2, pp. 229–246, Jun. 2012.
[11] M. Thuzar, "Urbanization in SouthEast Asia: Developing Smart Cities for the Future," *Reg. Outlook*, pp. 96–100, 2011.
[12] A. Caragliu, C. Del Bo, and P. Nijkamp, "Smart Cities in Europe," *J. Urban Technol.*, vol. 18, no. 2, pp. 65–82, Apr. 2011.
[13] T. Nam and T. A. Pardo, "Conceptualizing Smart City with Dimensions of Technology, People, and Institutions," in *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 2011, pp. 282–291.
[14] S. Zygiaris, "Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems," *J. Knowl. Econ.*, vol. 4, no. 2, pp. 217–231, 2013.
[15] A. F. Batisha, "Feasibility and sustainability of fog harvesting," *Sustain. Water Qual. Ecol.*, vol. 6, pp. 1–10, 2015.
[16] L. Junguo *et al.*, "Water scarcity assessments in the past, present, and future," *Earth's Futur.*, vol. 5, no. 6, pp. 545–559, Mar. 2017.
[17] F. R. Rijsberman, "Water scarcity: Fact or fiction?," *Agric. Water Manag.*, vol. 80, no. 1, pp. 5–22, 2006.
[18] J. Schewe *et al.*, "Multimodel assessment of water scarcity under climate change," *Proc. Natl. Acad. Sci.*, vol. 111, no. 9, p. 3245 LP-3250, Mar. 2014.
[19] N. W. Arnell, "Climate change and global water resources:SRES emissions and socio-economic scenarios," *Glob. Environ. Chang.*, vol. 14, no. 1, pp. 31–52, 2004.
[20] T. Oki and S. Kanae, "Global Hydrological Cycles and World Water Resources," *Science (80-.).*, vol. 313, no. 5790, p. 1068 LP-1072, Aug. 2006.

[21] J. ALCAMO, M. FLÖRKE, and M. MÄRKER, "Future long-term changes in global water resources driven by socio-economic and climatic changes," *Hydrol. Sci. J.*, vol. 52, no. 2, pp. 247–275, Apr. 2007.

[22] A. Hayashi, K. Akimoto, F. Sano, S. Mori, and T. Tomoda, "Evaluation of global warming impacts for different levels of stabilization as a step toward determination of the long-term stabilization target," *Clim. Change*, vol. 98, no. 1, p. 87, 2009.

[23] N. W. Arnell and S. N. Gosling, "The impacts of climate change on river flow regimes at the global scale," *J. Hydrol.*, vol. 486, pp. 351–364, 2013.

[24] S. N. Gosling and N. W. Arnell, "A global assessment of the impact of climate change on water scarcity," *Clim. Change*, vol. 134, no. 3, pp. 371–385, 2016.

[25] C. J. Vörösmarty, P. Green, J. Salisbury, and R. B. Lammers, "Global Water Resources: Vulnerability from Climate Change and Population Growth," *Science (80-.).*, vol. 289, no. 5477, p. 284 LP-288, Jul. 2000.

[26] S. Dolnicar and A. Hurlimann, "Drinking water from alternative water sources: differences in beliefs, social norms and factors of perceived behavioural control across eight Australian locations," *Water Sci. Technol.*, vol. 60, no. 6, p. 1433 LP-1444, Sep. 2009.

[27] S. Piao *et al.*, "The impacts of climate change on water resources and agriculture in China," *Nature*, vol. 467, p. 43, Sep. 2010.

[28] R. G. Raluy, L. Serra, and J. Uche, "Life Cycle Assessment of Water Production Technologies - Part 1: Life Cycle Assessment of Different Commercial Desalination Technologies (MSF, MED, RO) (9 pp)," *Int. J. Life Cycle Assess.*, vol. 10, no. 4, pp. 285–293, 2005.

[29] R. G. Raluy, L. Serra, J. Uche, and A. Valero, "Life Cycle Assessment of Water Production Technologies - Part 2: Reverse Osmosis Desalination versus the Ebro River Water Transfer (9 pp)," *Int. J. Life Cycle Assess.*, vol. 10, no. 5, pp. 346–354, 2005.

[30] J. Stokes and A. Horvath, "Life Cycle Energy Assessment of Alternative Water Supply Systems (9 pp)," *Int. J. Life Cycle Assess.*, vol. 11, no. 5, pp. 335–343, 2006.

[31] O. Klemm *et al.*, "Fog as a Fresh-Water Resource: Overview and Perspectives," *Ambio*, vol. 41, no. 3, pp. 221–234, 2012.

[32] M. Fessehaye, S. A. Abdul-Wahab, M. J. Savage, T. Kohler, T. Gherezghiher, and H. Hurni, "Fog-water collection for community use," *Renew. Sustain. Energy Rev.*, vol. 29, pp. 52–62, 2014.

[33] V. J. Camilo, T. Conrado, and B. D. D., "Fog interception by non-vascular epiphytes in tropical montane cloud forests: dependencies on gauge type and meteorological conditions," *Hydrol. Process.*, vol. 22, no. 14, pp. 2484–2492, Jun. 2008.

[34] R. J. Pilié, E. J. Mack, C. W. Rogers, U. Katz, and W. C. Kocmond, "The Formation of Marine Fog and the Development of Fog-Stratus Systems along the California Coast," *J. Appl. Meteorol.*, vol. 18, no. 10, pp. 1275–1286, Oct. 1979.

[35] S. B. LA, E. Werner, and B. Reto, “Fog as a Hydrologic Input,” *Encyclopedia of Hydrological Sciences*. 15-Apr-2006.
[36] S. Prada, M. Menezes de Sequeira, C. Figueira, and M. O. da Silva, “Fog precipitation and rainfall interception in the natural forests of Madeira Island (Portugal),” *Agric. For. Meteorol.*, vol. 149, no. 6, pp. 1179–1187, 2009.
[37] J. Cavelier and G. Goldstein, “Mist and fog interception in elfin cloud forests in Colombia and Venezuela,” *J. Trop. Ecol.*, vol. 5, no. 03, pp. 309–322, 1989.
[38] L. B. Hutley, D. Doley, D. J. Yates, and A. Boonsaner, “Water Balance of an Australian Subtropical Rainforest at Altitude: the Ecological and Physiological Significance of Intercepted Cloud and Fog,” *Aust. J. Bot.*, vol. 45, no. 2, p. 311, 1997.
[39] D. McJannet, J. Wallace, and P. Reddell, “Precipitation interception in Australian tropical rainforests: 2. Altitudinal gradients of cloud interception, stemflow, throughfall and interception,” *Hydrol. Process.*, vol. 21, no. 13, pp. 1703–1718, 2007.
[40] F. Holwerda, R. Burkard, W. Eugster, F. N. Scatena, A. G. C. A. Meesters, and L. A. Bruijnzeel, “Estimating fog deposition at a Puerto Rican elfin cloud forest site: comparison of the water budget and eddy covariance methods,” *Hydrol. Process.*, vol. 20, no. 13, pp. 2669–2692, 2006.
[41] L. A. Bruijnzeel, M. Mulligan, and F. N. Scatena, “Hydrometeorology of tropical montane cloud forests: emerging patterns: Hydrometeorology of tropical montane cloud forests,” *Hydrol. Process.*, vol. 25, no. 3, pp. 465–498, 2011.
[42] Y. Uehara and A. Kume, “Canopy Rainfall Interception and Fog Capture by Pinus pumila Regal at Mt. Tateyama in the Northern Japan Alps, Japan,” *Arctic¸ Antarct. Alp. Res.*, vol. 44, no. 1, pp. 143–150, 2012.
[43] C. Domínguez *et al.*, “Quantification of cloud water interception in the canopy vegetation from fog gauge measurements,” *Hydrol. Process.*, vol. 31, no. 18, pp. 3191–3205, Aug. 2017.
[44] P. Cereceda, R. S. Schemenauer, and M. Suit, “An alternative water supply for chilean coastal desert villages,” *Int. J. Water Resour. Dev.*, vol. 8, no. 1, pp. 53–59, Mar. 1992.
[45] R. S. Schemenauer and P. Cereceda, “A Proposed Standard Fog Collector for Use in High-Elevation Regions,” *J. Appl. Meteorol.*, vol. 33, no. 11, pp. 1313–1322, Nov. 1994.
[46] E. C. David and R. W. Zimmerman, “Compressibility and shear compliance of spheroidal pores: Exact derivation via the Eshelby tensor, and asymptotic expressions in limiting cases,” *Int. J. Solids Struct.*, vol. 48, no. 5, pp. 680–686, 2011.
[47] E. S. Shanyengana, J. R. Henschel, M. K. Seely, and R. D. Sanderson, “Exploring fog as a supplementary water source in Namibia,” *Atmos. Res.*, vol. 64, no. 1, pp. 251–259, 2002.
[48] M. V. Marzol Jaén, “Fog water collection in a rural park in the Canary Islands (Spain),” *Atmos. Res.*, vol. 64, no. 1, pp. 239–250, 2002.

[49] J. Olivier, "Fog harvesting: An alternative source of water supply on the West Coast of South Africa," *GeoJournal*, vol. 61, no. 2, p. 203, 2004.

[50] S. A. Abdul-Wahab, A. M. Al-Damkhi, H. Al-Hinai, K. A. Al-Najar, and M. S. Al-Kalbani, "Total fog and rainwater collection in the Dhofar region of the Sultanate of Oman during the monsoon season," *Water Int.*, vol. 35, no. 1, pp. 100–109, Feb. 2010.

[51] M. J. Estrela, J. A. Valiente, D. Corell, and M. M. Millán, "Fog collection in the western Mediterranean basin (Valencia region, Spain)," *Atmos. Res.*, vol. 87, no. 3, pp. 324–337, 2008.

[52] A. Ritter, M. C. Regalado, and C. J. Guerra, "Quantification of Fog Water Collection in Three Locations of Tenerife (Canary Islands)," *Water* , vol. 7, no. 7. 2015.

[53] K. A. Brauman, D. L. Freyberg, and G. C. Daily, "Forest structure influences on rainfall partitioning and cloud interception: A comparison of native forest sites in Kona, Hawai'i," *Agric. For. Meteorol.*, vol. 150, no. 2, pp. 265–275, 2010.

[54] A. Ritter, C. M. Regalado, and G. Aschan, "Fog Water Collection in a Subtropical Elfin Laurel Forest of the Garajonay National Park (Canary Islands): A Combined Approach Using Artificial Fog Catchers and a Physically Based Impaction Model," *J. Hydrometeorol.*, vol. 9, no. 5, pp. 920–935, 2008.

[55] E. Gioda, G. R. Espejo, and A. Acosta Baladón, "Fog collectors in tropical areas," in *Proceedings of the symposium on precipitation and evaporation*, 1993, pp. 4–20.

[56] E. P. C. Juvik J.O., "A climatology of mountain fog on Mauna Loa, Hawai'i Island," *Univ. Hawai'i Water Resour. Cent. Tech. Rep.*, no. 118, 1978.

[57] K. F. A. Frumau, R. Burkard, S. Schmid, L. A. (Sampurno) Bruijnzeel, C. Tobón, and J. C. Calvo-Alvarado, "A comparison of the performance of three types of passive fog gauges under conditions of wind-driven fog and precipitation," *Hydrol. Process.*, vol. 25, no. 3, pp. 374–383, 2011.

[58] J. Van Heerden, J. Olivier, and L. van Schalkwyk, "Van Heerden, J., Olivier, J., van Schalkwyk, L., 2010. Fog water systems in South Africa: an update. In: Proceedings of the 5th International Conference on Fog, Fog Collection and Dew, July 25–30, 2010, Münster, p. 160.," *In: Proceedings of the 5th International Conference on Fog, Fog Collection and Dew*, 2010, p. 160.

[59] P. Cereceda, P. Hernández, J. Leiva, and J. Rivera, "Agua de Niebla," Santiago: Pontificia Universidad Católica, 2014.

[60] K. J. Tiedemann and A. Lummerich, "Fog harvesting on the verge of economic competitiveness," in *In: Proceedings of the 5th International Conference on Fog, Fog Collection and Dew*, 2010, p. 192.

[61] C. M. Regalado and A. Ritter, "The performance of three fog gauges under field conditions and its relationship with meteorological variables in an exposed site in Tenerife (Canary Islands)," *Agric. For. Meteorol.*, vol. 233, pp. 80–91, 2017.

[62] C. Tomasi and F. Tampieri, "Features of the proportionality coefficient in the relationship between visibility and liquid water content in haze and fog," *Atmosphere (Basel).*, vol. 14, no. 2, pp. 61–76, Jun. 1976.
[63] K. F. A. Frumau, L. A. Bruijnzeel, C. Tobón, and others, "Hydrological measurement protocol for montane cloud forests. Annex 2, Final Technical Report on DFID-FRP Project no. R7991.," 2006.
[64] R. Gunn and G. D. Kinzer, "The terminal velocity of fall for water droplets in stagnant air," *J. Meteorol.*, vol. 6, no. 4, pp. 243–248, 1949.
[65] S. R. Herwitz, "Interception storage capacities of tropical rainforest canopy trees," *J. Hydrol.*, vol. 77, no. 1–4, pp. 237–252, 1985.
[66] INEC, "Resultados del censo 2010 de población y vivienda en el Ecuador," Quito, Ecuador, 2010.
[67] E. C. and T. A., "The influence of the Galápagos Islands on tropical temperatures, currents and the generation of tropical instability waves," *Geophys. Res. Lett.*, vol. 31, no. 15, Aug. 2004.
[68] M. Trueman and N. D'Ozouville, "Characterizing the Galapagos terrestrial climate in the face of global climate change," *Galapagos Res.*, vol. 67, pp. 26–37, 2010.
[69] A. Pryet *et al.*, "Quantification of cloud water interception along the windward slope of Santa Cruz Island, Galapagos (Ecuador)," *Agric. For. Meteorol.*, vol. 161, pp. 94–106, 2012.
[70] R. B. Dunbar, G. M. Wellington, M. W. Colgan, and P. W. Glynn, "Eastern Pacific sea surface temperature since 1600 A.D.: The record of climate variability in Galápagos Corals," *Paleoceanography*, vol. 9, no. 2, pp. 291–315, 1994.
[71] C. G. Domínguez *et al.*, "Comparison of deep percolation rates below contrasting land covers with a joint canopy and soil model," *J. Hydrol.*, vol. 532, pp. 65–79, 2016.
[72] D. J. Geist, A. R. McBIRNEY, and R. A. Duncan, "Geology and petrogenesis of lavas from San Cristobal Island, Galapagos Archipelago," *Geol. Soc. Am. Bull.*, vol. 97, no. 5, p. 555, 1986.
[73] M. Adelinet, J. Fortin, N. D'Ozouville, and S. Violette, "The relationship between hydrodynamic properties and weathering of soils derived from volcanic rocks, Galapagos Islands (Ecuador)," *Environ. Geol.*, vol. 56, no. 1, pp. 45–58, 2008.
[74] A. Pryet, N. d'Ozouville, S. Violette, B. Deffontaines, and E. Auken, "Hydrogeological settings of a volcanic island (San Cristóbal, Galapagos) from joint interpretation of airborne electromagnetics and geomorphological observations," *Hydrol. Earth Syst. Sci.*, vol. 16, no. 12, pp. 4571–4579, 2012.
[75] S. Violette, N. d'Ozouville, A. Pryet, B. Deffontaines, J. Fortin, and M. Adelinet, "Hydrogeology of the Galápagos Archipelago: An Integrated and Comparative Approach Between Islands," in *Geophysical Monograph Series*, K. S. Harpp, E. Mittelstaedt, N. d'Ozouville, and D. W. Graham, Eds. Hoboken, New Jersey: John Wiley & Sons, Inc, 2014, pp. 167–183.

[76] N. D'Ozouville *et al.*, "Extensive perched aquifer and structural implications revealed by 3D resistivity mapping in a Galapagos volcano," *Earth Planet. Sci. Lett.*, vol. 269, no. 3–4, pp. 518–522, May 2008.
[77] C. Domínguez, "Integrated Hydrogeological Study of San Cristobal Island (Galapagos)," Université Pierre et Marie Curie - Paris VI, Paris, 2016.
[78] M. Adelinet, "Etude du fonctionnement hydrologique des bassins versant de San Cristobal (Galapagos - Equateur," Université Pierre et Marie Curie - Paris {VI}, Paris, 2005.
[79] N. d'Ozouville, "Etude du Fonctionnement Hydrologique Dans les Iles Galápagos : caractérisation d'un milieu volcanique insulaire et préalable à la gestion de la ressource.," 2007.
[80] (CGREG) Consejo de Gobierno del Régimen Especial de Galápagos, "Censo de Unidades de Producción Agropecuaria de Galápagos," 2014.
[81] D. and R. (CISPDR) Changjiang Institute of Survey, Planning, "La planificación de los recursos hídricos de las islas galápagos," 2015.
[82] R. S. Schemenauer and P. I. Joe, "The collection efficiency of a massive fog collector," *Atmos. Res.*, vol. 24, no. 1, pp. 53–69, 1989.
[83] J. C. Villegas, C. Tobón, and D. D. Breshears, "Fog interception by non-vascular epiphytes in tropical montane cloud forests: dependencies on gauge type and meteorological conditions," *Hydrol. Process.*, vol. 22, no. 14, pp. 2484–2492, 2008.
[84] W. Eugster, R. Burkard, F. Holwerda, F. N. Scatena, and L. A. S. Bruijnzeel, "Characteristics of fog and fogwater fluxes in a Puerto Rican elfin cloud forest," *Agric. For. Meteorol.*, vol. 139, no. 3, pp. 288–306, 2006.
[85] T. W. Giambelluca and D. Nullet, "Influence of the trade-wind inversion on the climate of a leeward mountain slope in {Hawaii}," *Clim. Res.*, vol. 1, no. 3, pp. 207–216, 1991.
[86] A. Thompson, "Simulating the adiabatic ascent of atmospheric air parcels using the cloud chamber," *Dep. Meteorol. Penn State*, 2007.
[87] J. de D. Rivera, "Aerodynamic collection efficiency of fog water collectors," *Atmos. Res.*, vol. 102, no. 3, pp. 335–342, 2011.
[88] C. M. Regalado and A. Ritter, "The design of an optimal fog water collector: A theoretical analysis," *Atmos. Res.*, vol. 178–179, pp. 45–54, 2016.
[89] D. M. and T. Fernandez Alicia and Weiss-Penzias, Peter S. and Zhang, Bong June and Sorensen, Deckard and Cohen, Robert E. and McKinley, Gareth H. and Kleingartner, Justin and Oliphant, Andrew and Bowman, Matthew, "Fog Water Collection Effectiveness: Mesh Intercomparisons," *Aerosol Air Qual. Res.*, vol. 18, no. 1, pp. 270–283, 2018.
[90] F. M. White, *Fluid Mechanics*, Second. New York: McGraw-Hill Book Company, 1986.
[91] H. W. Vogelmann, "Fog Precipitation in the Cloud Forests of Eastern Mexico," *Bioscience*, vol. 23, no. 2, pp. 96–100, Feb. 1973.
[92] G. A. Al-hassan, "Fog Water Collection Evaluation in Asir Region–Saudi Arabia," *Water Resour. Manag.*, vol. 23, no. 13, pp. 2805–2813, 2009.

[93] Z. Mileta, M., Likso, "Fog water collection with SFC on the mountain Velebit (Croatia) during the period 2000–2009," in *Proceedings of the 5th International Conference on Fog, Fog Collection and Dew, July 25–30, 2010*, 2010, pp. 240–242.
[94] J. Olivier, "Fog-water harvesting along the West Coast of South Africa: A feasibility study," *Water SA*, vol. 28, no. 4, pp. 349–360, 2002.
[95] A. Sabino, "Fog collection in the natural park of Serra Malagueta. An alternative source of water for the communities.," in *Proceedings of the 4th international conference on fog, fog collection and dew*, 2007, pp. 425–428.
[96] P. Echeverría, "potencial de la captura de niebla para uso doméstico rural y riego durante la época seca del año en la isla San Cristobal, Galápagos, Ecuador," Escuela Politécnica Nacional, Ecuador, 2018.
[97] M. Rajaram, X. Heng, M. Oza, and C. Luo, "Enhancement of fog-collection efficiency of a Raschel mesh using surface coatings and local geometric changes," *Colloids Surfaces A Physicochem. Eng. Asp.*, vol. 508, pp. 218–229, 2016.

Sobre los autores



Christian Domínguez. (Y'1987-M'07) recibió el título de Ingeniero Civil de la Escuela Politécnica Nacional (2011) y el título de doctor en geociencia, recursos naturales y medio ambiente de la Universidad Pierre et Marie Curie (2016). Actualmente es profesor titular de la Carrera de Ingeniería Civil de la Universidad Politécnica Salesiana.



Paola Alejandra Echeverría Garcés (Y'1991-M'09). Paola Echeverría es Ingeniera Ambiental, obtuvo su título en la Facultad de Ingeniería Civil y Ambiental de la Escuela Politécnica Nacional en el año 2018.



Marcos Villacís (Y'1975 – M'06) recibió el título de Ingeniero Civil de la Escuela Politécnica Nacional (2001), realizó una maestría en Hidrología e Hidrogeología en la Universidad Pierre et Marie Curie (2003) y tiene un PhD en Aguas Continentales y Sociedad en la Universidad Montpellier 2 (2008). Actualmente es profesor titular del departamento de Ingeniería Civil y Ambiental de la Escuela Politécnica Nacional (EPN). En el ámbito de gestión académica, fue decano

(2015-2016) y jefe de departamento (2016-2018) de la Facultad Ingeniería Civil y Ambiental de la EPN. Su investigación se enfoca principalmente en la hidro-glaciología y el cambio climático, siendo co-director del proyecto de investigación "Linking Gobal Change with soil an water conservation in the high Andes" y co-director del Laboratorio Mixto Internacional GREATICE (Glaciares y Recursos Hídricos de Altura: Indicadores de Cambios Ambientales). En la actualidad, es editor invitado de la Revista Politécnica de la EPN y de la Revista Científica La Granja de la UPS.



Sophie Violette (Y'1966) recibió su PhD (1993) y luego un HDR (2003) en Hidrogeología de la Universidad Pierre et Marie Curie (UPMC). Actualmente es investigadora en el Laboratorio de Geología UMR.8538 de la Escuela Normal Superior (Francia) y es profesora titular en Hidrogeología en la Universidad Sorbona desde 1995. Sophie trabajó como científica visitante en la Universidad John Hopkins dentro del equipo de Prof. Grant Garven (1999) y en la Universidad Stanford dentro del equipo del Prof. Steve Gorelick (2005). El principal objetivo de sus actividades de investigación es entender los flujos de aguas subterráneas a varias escalas espaciales y temporales a partir de la identificación de: geometría en 3D (datos geofísicos y geológicos), propiedades (físicas, mecánicas, hidrodinámicas, transferencia de solutos y de calor), procesos físicos envueltos y condiciones de borde (tectónica, clima y el bombeo de aguas subterráneas). Sus actividades científicas incluyen una larga colección de modelos de flujos hidrológicos y tiene antecedentes multidisciplinarios combinados con experiencia en campo en regiones con varios climas como: Francia (islas volcánicas y cuencas sedimentarias), Bulgaria y Madagascar (lechos rocosos), India (cuencas sedimentarias), Ecuador e Indonesia (volcanes basálticos y andesíticos). Su experticia ha sido fortalecida por su participación y dirección de proyectos de investigación con colaboradores nacionales e internaciones. Su investigación ha sido financiada por: organismos públicos (AESN, ANDRA, ANR, ANRT, Chancellery of Universities, CRIF, INSU, IRSN, SEDIF), compañías privadas (DANONE, EDF, ESA, GDF-Suez), fundaciones (Ensemble, Foundation of France, Schlumberger, Véolia) y organismos internacionales (PNUD, World bank). Ella ha servido como experta en hidrogeología en varios comités internaciones (ANDRA-COS, IFPEN-CS, Universities-CSE, INRA-CSS).

CAPÍTULO IV

Estudio de propagación de ondas sísmicas – caso particular de sitio

Study of seismic waves propagation - site specific case

José Alviar*
Universidad Politécnica Salesiana Ecuador
*Email: jalviar@ups.edu.ec

Resumen

Uno de los problemas con que se enfrenta el ingeniero civil es el diseño adecuado de estructuras. En ese sentido, existen normas a nivel mundial que regulan el diseño y construcción de edificaciones solicitadas por acciones sísmicas. El objeto es hacer a las estructuras dúctiles, es decir que puedan experimentar grandes deformaciones, sin pérdida de resistencia y que no colapsen aun cuando actúe el sismo más severo.
En numerosas normas a nivel mundial se establece realizar estudios particulares de sitio o de propagación de ondas sísmicas en obras importantes, sin embargo no se descarta su aplicación en obras que pudiesen ser consideradas como poco importantes por el aspecto económico, sin dejar de serlo por el aspecto de salvar vidas. En ese sentido, en aquellos casos en que la edificación se encuentre en la cercanía a fallas activas o exista la presencia de estratos de suelos susceptibles a licuar o de suelos que degradan su rigidez o resistencia con la deformación bajo el efecto de carga cíclica deben ser objeto de la realización de estudios particulares de sitio. Por ello, se presenta en el siguiente trabajo la metodología utilizada para realizar un estudio particular de sitio, mostrando entre otros aspectos la degradación de rigidez y resistencia de los suelos, la evaluación del potencial de licuación de arenas, los niveles de deformación que experimentan los estratos de suelos y el espectro de diseño obtenido como insumo para el análisis dinámico de estructuras.

Palabras clave

Deformación cíclica, efectos de sitio, espectro de diseño, licuación, ondas sísmicas, propagación, respuesta dinámica.

Abstract

One of the problems facing the civil engineer is the proper design of structures. In that sense, there are worldwide standards that regulate the design and construction of buildings requested by seismic actions. The object is to make the structures ductile, so, it can undergo great deformations, without loss of resistance and do not collapse even when the most severe earthquake acts.
In numerous standards at world-wide it is established to realize particular studies of site or to evaluate the propagation of seismic waves in major constructions, nevertheless it is not discarded its application in works that could be considered as unimportant by the economic aspect, without ceasing to be for the aspect of saving lives. In this sense, in those cases in which the building is in the vicinity of active faults or there is the presence of strata of soils susceptible to liquefying or of soils that degrade their rigidity or resistance with the deformation under the effect of cyclic loading must be object of the realization of particular site studies. For this reason, the methodology used to carry out a particular site study is presented in the following work, showing, among other aspects, the degradation of soil stiffness and resistance, the evaluation of the liquefaction potential of sands, the deformation levels reached by the soil strata and the design spectrum obtained as input for the dynamic analysis of structures.

Keywords

Cyclic deformation, design spectrum, dynamic response, liquefaction, propagation, seismic waves, site effects.

1. Introducción

Existen en el país numerosos sistemas de fallas activas que por su cercanía a obras civiles, bien sea por construir, u obras existentes donde se desee evaluar cuál es el comportamiento dinámico de la estructura y qué medidas de reforzamiento deben ser implementados ante acciones sísmicas, amerita la realización de estudios particulares de sitio. Igualmente, la presencia de estratos o depósitos de suelos potencialmente licuables [1-2-3] o de suelos que degradan su rigidez o resistencia bajo el efecto sísmico deben ser objeto de un estudio particular de sitio, a fin de diseñar apropiadamente el sistema de cimentación, así como el cuantificar los niveles de degradación de rigidez y resistencia que puedan experimentar los suelos, determinar el factor de seguridad contra la licuación de los depósitos granulares sueltos saturados y la generación del espectro de diseño que servirá como insumo para la realización del análisis

dinámico de la estructura. En ese sentido se presenta a continuación la metodología seguida para realizar un estudio particular de sitio mostrando su aplicación en el análisis de licuación de los suelos y la generación del espectro de diseño para un sitio en particular.

2. Materiales y métodos

El presente trabajo constituye una investigación cuantitativa del tipo cuasi-experimental [4], por cuanto se estudia la relación causa-efecto de fenómenos naturales como los sismos, sobre depósitos de suelos donde no existe un control riguroso de la variable independiente que es la historia de aceleraciones sísmicas. Los mismos provienen de registros instrumentales disponibles en la red sismológica mundial [5]. De acuerdo con la definición operacional de variables, en este estudio se presentan:

Variables Independientes: comprende los registros sísmicos medidos en roca y los perfiles litológicos caracterizados por su espesor, peso unitario y velocidad de ondas de corte, entre otros.

Variables Intervinientes: comprende la herramienta o algoritmo de cálculo [6] que permite cuantificar los efectos que produce el sismo en el comportamiento dinámico del perfil del suelo. Esta variable no será sometida a investigación, por cuanto se parte de la premisa de que el algoritmo de cálculo ha sido desarrollado y validado por personal altamente especializado [7-8].

Variables Dependientes: comprende la historia de aceleraciones en superficie, espectros de respuesta, lazos de histéresis, trayectoria de esfuerzos, incremento en presión de poros, entre otros.

La muestra de estudio se circunscribe a depósitos de suelos, saturados, susceptibles a experimentar el fenómeno de licuación, estratos de suelos que pueden degradar su rigidez y resistencia bajo el efecto de carga cíclica, estratos de arenas o arcillas competentes en los que se desea cuantificar los niveles de deformación cíclico, así como los niveles de degradación que pueden alcanzar a fin de elaborar las curvas de capacidad portante de cimentaciones profundas con profundidad.

La metodología a seguir para cumplir con el objetivo propuesto en este estudio se presenta continuación:

Recopilación bibliográfica de trabajos realizados que tienen relación con el tema objeto de estudio, destacándose entre ellos: evaluación del potencial de licuación [9-10], modelos constitutivos que consideren degradación de resistencia, cambio en rigidez y amortiguamiento en los suelos con la deformación y aumento de presión de poros [11-12-13-14], entre otros.

Ubicación y selección de código de cálculo o programa que permita evaluar la respuesta dinámica de suelos, y particularmente considere los cambios de rigidez y amortiguamiento con el aumento de presión de poros, así como su disipación en el tiempo.

Construir un perfil de suelo representativo del lugar en donde se realizará el estudio particular de sitio. Seleccionar sismos instrumentales registrados en roca en correspondencia con el nivel de amenaza esperado en el sitio de estudio y su distancia o cercanía a las fallas activas. Aplicar en la base del perfil geotécnico los acelerogramas de entrada definidos previamente.

Obtener la historia de aceleraciones en superficie, lazos de histéresis, historia de aumento de presión de poros, entre otros.

A partir de la historia de aceleraciones en superficie construir los espectros de respuesta para una estructura elástica de 1 grado de libertad (gdl).

Realizar el análisis de las respuestas obtenidas cuantificando la aceleración máxima en superficie, la amplificación y la generación de espectros de respuesta. Así mismo cuantificar los niveles de deformación que se producen a lo alto del depósito, la determinación de los niveles de degradación de rigidez y resistencia de los suelos que componen la columna litológica, la evaluación de los factores de seguridad contra la licuación que puedan sufrir los suelos granulares, sueltos, saturados.

2.1. *Respuesta dinámica de suelos bajo acciones sísmicas*

El estudio particular de sitio o la respuesta dinámica del suelo bajo acciones sísmicas consiste en imponer en la base del depósito sismos representativos, cuyas características se correspondan con el estudio de amenaza, en cuanto a magnitud esperada, distancia epicentral y aceleración en roca, los cuales se propagan a través del

depósito de suelo o columna litológica y permiten determinar los valores de esfuerzos de corte y deformaciones máximas impuestas por el sismo, con profundidad, y las aceleraciones que podrían producirse o el espectro de respuesta.

Una vez determinadas las deformaciones máximas de corte con profundidad o la movilidad cíclica, podrán determinarse los valores de degradación de la rigidez y resistencia, utilizando resultados de ensayos de corte cíclicos normalizados reportados en la literatura [11] o de ensayos de corte cíclico a esfuerzo controlado o deformación controlada realizado en muestras representativas de la zona objeto de estudio.

Con los valores de esfuerzos de corte máximos con profundidad impuestos por el sismo se podrán determinar los valores de CSR o razón de esfuerzo cíclico correspondiente a 0.65 veces el valor máximo de corte dividido entre el esfuerzo vertical efectivo. En el caso de los suelos granulares se determinará si el estrato granular licuará o no al comparar los valores de CSR con los valores de CRR (cyclic resistance ratio) o la resistencia al corte cíclico de las arenas derivado de la resistencia de punta del cono del ensayo CPTU o del ensayo SPT [15]. En el caso de los suelos arcillosos, basado en los niveles de deformación alcanzados se degradará su rigidez y resistencia (movilidad cíclica).

Con los valores de rigidez y resistencia degradados con profundidad y la identificación de los estratos granulares que experimentan licuación, permitirá definir los factores que modifican tanto la rigidez como la resistencia de las curvas p-y y t-z al ser sometido el pilote o cimentación profunda a carga lateral o axial, respectivamente, tanto en el campo libre como en el campo cercano.

2.2. *Estudio de la respuesta dinámica de un sitio en particular*

Para la ejecución de un estudio particular de sitio se realizarían las siguientes actividades:

- Selección de sismos instrumentales de la red mundial, en lo posible, con origen tectónico similar al de la zona objeto de estudio.
- Selección de los perfiles geotécnicos y sus parámetros para realizar las corridas con el programa CYCLIC 1D.

- Procesamiento de la salida de resultados y elaboración de perfiles de respuesta de las deformaciones máximas de corte y esfuerzos de corte máximos con profundidad.
- Elaboración de espectros de respuesta en superficie para cada uno de los sismos incluidos en el estudio.
- Determinación de la mediana de los espectros y el correspondiente a la media más una desviación estándar.
- Elaboración de los espectros de diseño.

Se debe tener especial cuidado de seleccionar sismos cuyos registros han sido medidos en roca, sin estar influenciados por la presencia de estructuras cercanas. Este tipo de registros son considerados como obtenidos en afloramientos rocosos presentes en el campo libre (freefield-outcropping).

La cantidad de sismos seleccionados deberán ser como mínimo 3 en correspondencia con la norma ecuatoriana NEC-SE-DS [16], los cuales se consideran representativos y cumplen con lo mínimo exigido en las normativas vigentes para la realización de estudios particulares de sitio. Para la realización del presente estudio se han seleccionado 7 sismos instrumentales para realizar los análisis.

3. Movilidad cíclica

Con el objeto de determinar los niveles de deformación inducidos por los sismos en los estratos que componen el Perfil Litológico, se recurrió al uso del programa o código de cálculo Cyclic 1-D.

El programa CYCLIC 1-D [6-7-8], fue desarrollado para evaluar la respuesta dinámica de suelos donde está involucrado el comportamiento no lineal e histerético asociado a los cambios en rigidez y amortiguamiento con la deformación por efecto sísmico, además de incluir los incrementos y disminuciones de la presión de poros y como ello afecta los esfuerzos totales y efectivos y en consecuencia la resistencia al corte del material.

Mediante el uso del programa se obtendrán los valores de deformación por corte máximos con profundidad a fin de identificar por sectores cuales son los niveles de deformación máximos alcanzados cuando la columna de suelo es sometida en su base al escenario de sismos previamente establecido de los estudios de amenaza.

A partir de los valores de deformaciones máximas alcanzadas (g max) se determinará la deformación por corte cíclica equivalente (0.65 g max), que al introducirlo en la Figura 1, envolvente de resultados obtenidos de ensayos de corte cíclicos realizados en muestras seleccionadas, se obtendrían los respectivos valores de degradación de rigidez que se producen en el campo libre.

Particularmente se presenta en la figura 1, en el eje de las ordenadas la degradación del módulo de corte (G/Gmax) con la deformación por corte cíclica al someter las muestras de suelos a ensayos de corte a deformación controlada bajo la acción de 15 a 30 ciclos de carga y descarga.

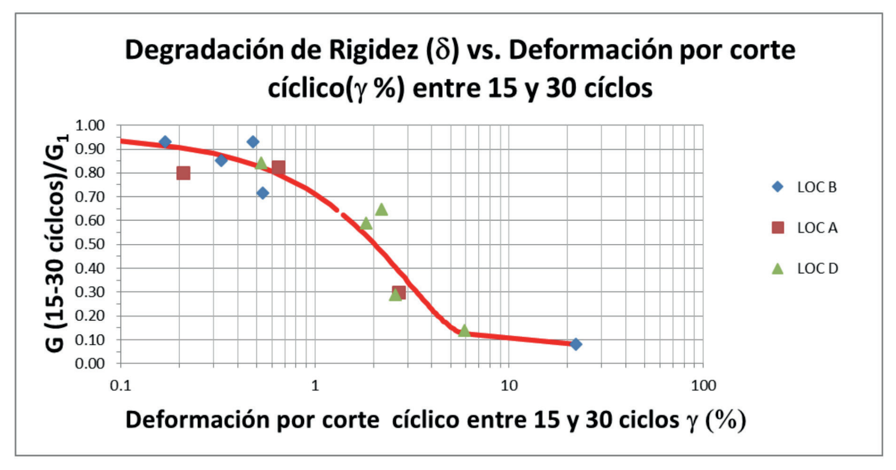


Figura 1. Degradación de rigidez de los suelos con la deformación por corte cíclico en muestras sometidas a ensayos de corte cíclico (DSS)

Los valores derivados de los ensayos dinámicos indican que hasta deformaciones por corte cíclicos de 0.6%, la degradación de rigidez (δ) es del orden de 0.8. Para deformaciones por corte de 1% y 2% el valor de δ es de 0.7 y 0.5 respectivamente.

El estudio de propagación de ondas sísmicas a través de la columna litológica del suelo ha permitido elaborar la gráfica de los valores de deformación equivalentes correspondientes a 0.65 veces la deformación máxima que se representan en las Figuras 2 y 3 al utilizar sismos representativos del estudio de amenaza considerando periodos de

retorno de 200 años, escenario SLE (Strength Level Earthquake) y 1000 años, escenario DLE (Ductility Level Earthquake) respectivamente.

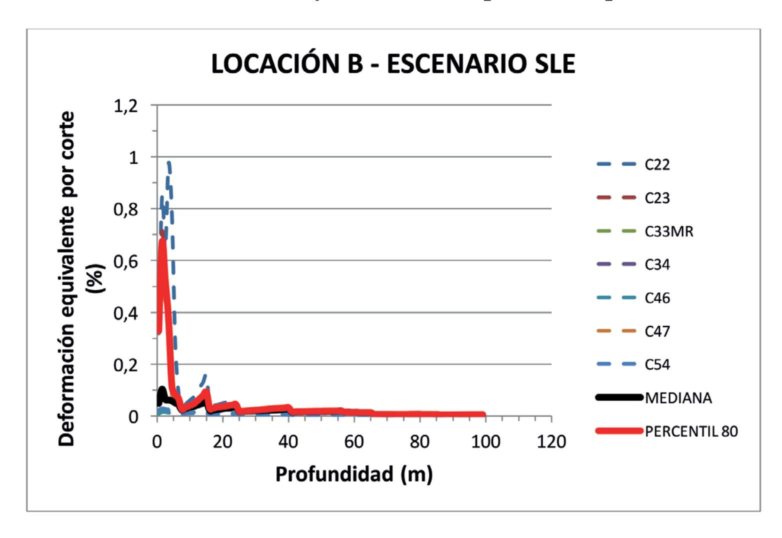


Figura 2. Deformación equivalente por corte cíclico con profundidad – Escenario SLE, periodo de retorno 200 años

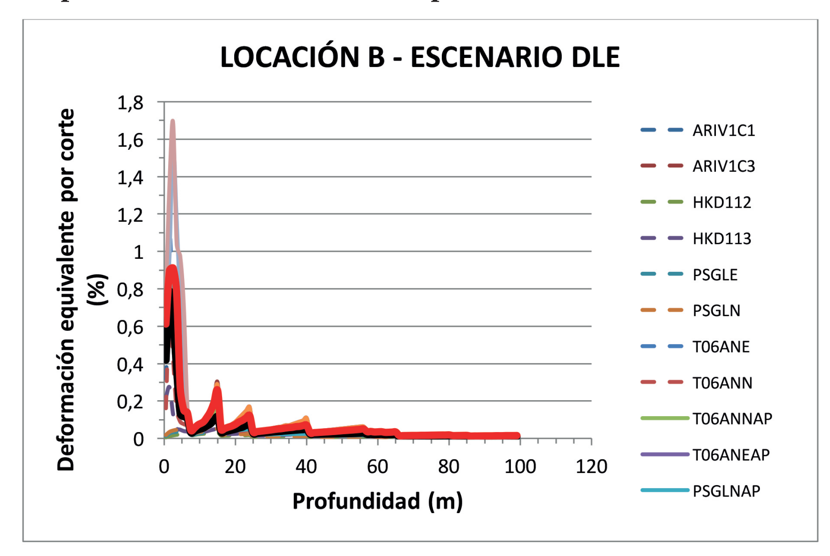


Figura 3. Deformación equivalente por corte cíclico con profundidad – escenario DLE, periodo de retorno 1000 años

De las figuras 2 y 3 se concluye que los estratos ubicados por debajo de los 7 m de profundidad experimentan una movilidad o una deformación por corte cíclica equivalente que no supera el 0.3% bajo los escenarios SLE y DLE, la cual va disminuyendo con profundidad, valor bajo de movilidad cíclica que producirá bajos valores de degradación de rigidez y resistencia. Mientras que el estrato de suelo ubicado entre los 0 y 7 m de profundidad, correspondiente a arenas sueltas saturadas experimenta altos niveles de deformación cuyo análisis del factor de seguridad contra la licuación se presenta en el siguiente aparte.

4 . Licuación de suelos

La licuación de los suelos es un fenómeno que ocurre en suelos granulares, sueltos, saturados ubicados por debajo del nivel freático. Se trata de arenas con poco contenido de finos o limos no plásticos que al ser sacudidos por la acción sísmica, tienden a disminuir de volumen. El agua intersticial presente en los poros o espacios vacíos no tiene la suficiente rapidez para migrar, lo cual provoca que la presión de poros aumente y se iguale al esfuerzo efectivo, bajo dicha condición se produce una pérdida significativa de resistencia haciendo que el suelo se comporte como un líquido y alcance lo que se conoce como licuación con un valor de resistencia al corte muy bajo o resistencia residual.

Una simulación realizada con el programa Cyclic 1-D, el cual permite aplicar en la base del depósito sismos instrumentales registrados en la red mundial permite visualizar como es el comportamiento bajo la acción sísmica.

El esfuerzo de corte impuesto por el sismo tiende a aumentar linealmente con profundidad, como se muestra en la Figura 4, al igual que el esfuerzo confinante, al cual está asociada una mayor resistencia al corte o esfuerzo cedente, pero en algún punto del perfil, el esfuerzo de corte impuesto por el sismo iguala al esfuerzo cedente, haciendo que el material alcance la falla y se produzcan los mayores valores de deformación, Figura 5. Simultáneamente ocurre el aumento en presión de poros que hace que los esfuerzos efectivos disminuyan, al igual que la resistencia al corte del material, y en caso de que se iguale el esfuerzo confinante con la presión de poros, el esfuerzo confinante efectivo se anula, produciéndose lo que se denomina licuación, Figura 6.

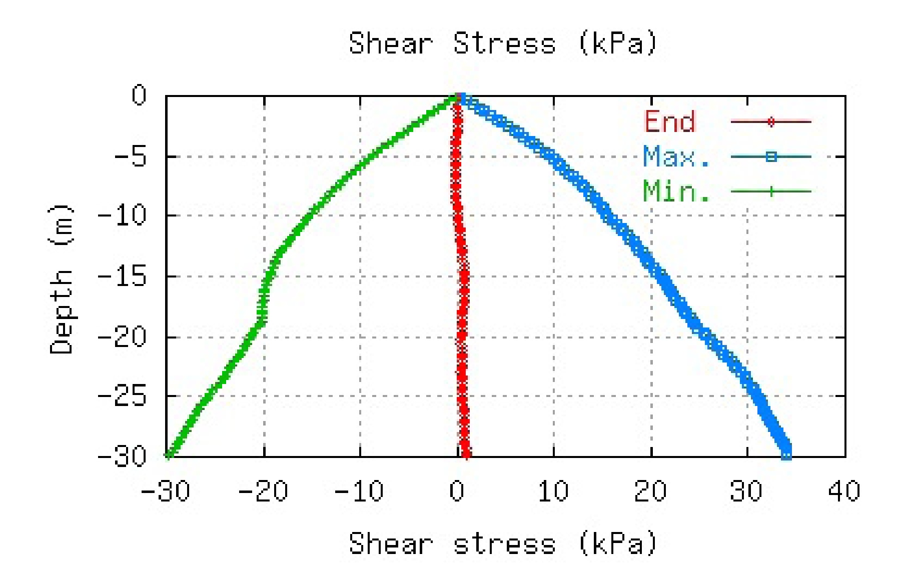


Figura 4. Esfuerzo de corte impuesto por el sismo vs profundidad

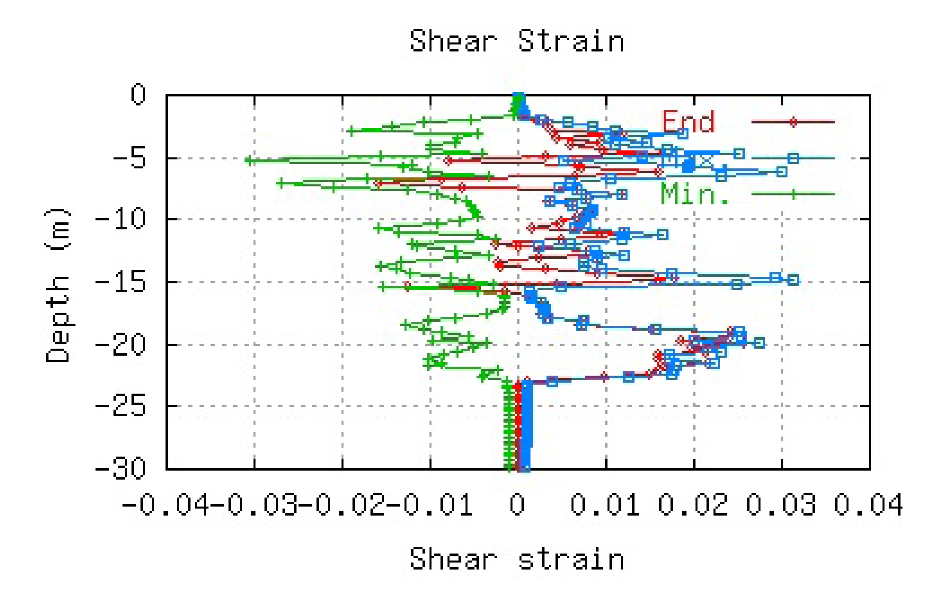


Figura 5. Deformaciones de corte vs profundidad

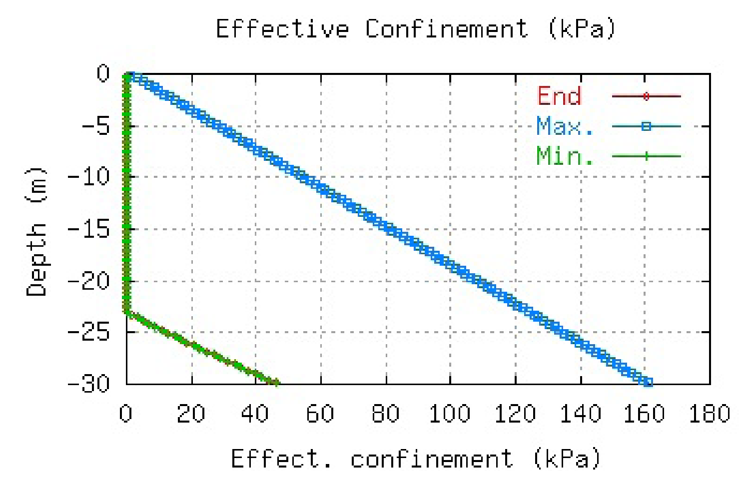


Figura 6. Esfuerzo de confinamiento efectivo vs profundidad

Los resultados anteriormente mostrados se corresponden con la salida obtenida con el programa Cyclic 1-D para un perfil de 30 m de espesor, al aplicar en la base del depósito el sismo ocurrido en Chile en el año 1985. En este caso, se produjo licuación inicial desde los 0 m hasta los 23 m, dado que el esfuerzo confinante efectivo se anuló para este rango de profundidades (Figura 6).

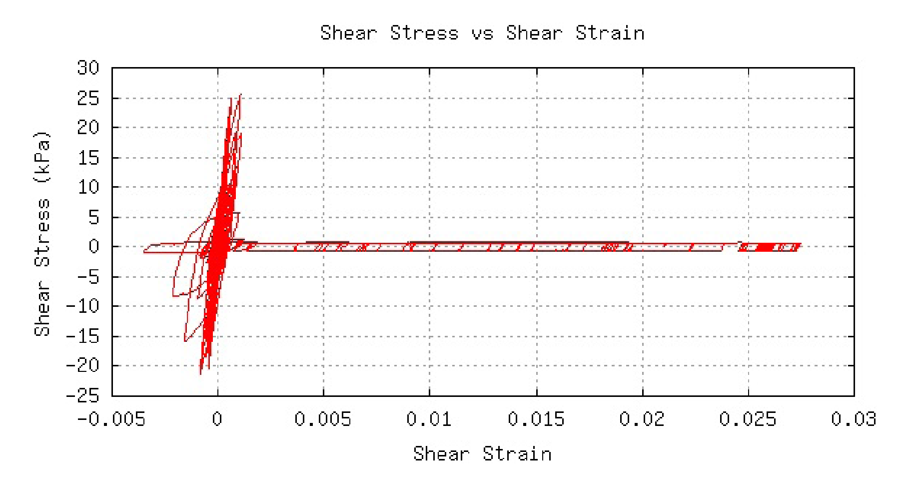


Figura 7. Lazo de histéresis a 20 m de profundidad

Para observar si se produjo licuación en el sentido estricto, es decir si el suelo experimentó grandes deformaciones bajo estado residual de resistencia al corte, con disminución de rigidez y escasa variabilidad en la presión de poros en el estado licuado (comportamiento contractivo), se presenta a continuación los lazos de histéresis a 20 m de profundidad (Figura 7), la historia de incrementos de presión de poros (Figura 8) y la trayectoria de esfuerzos correspondiente (Figura 9).

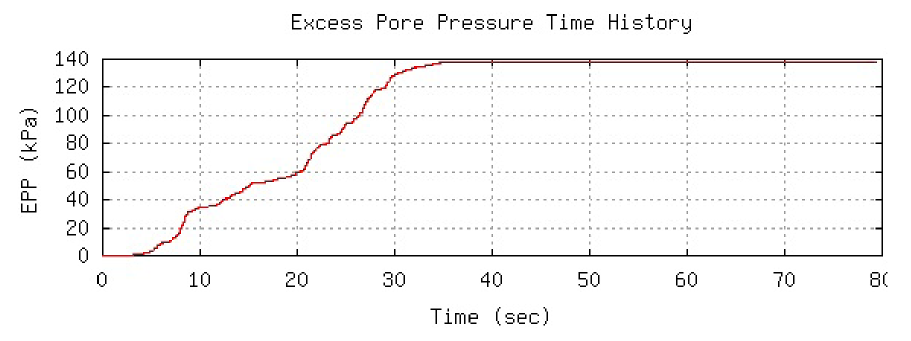


Figura 8. Aumento de presión de poros a 20 m de profundidad

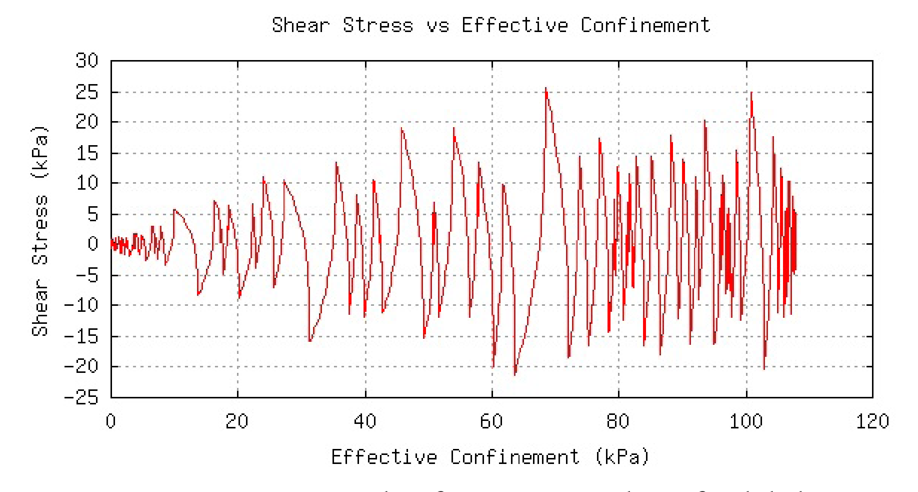


Figura 9. Trayectoria de esfuerzos a 20 m de profundidad

Se destaca de las figuras anteriores que después de ocurrida la licuación inicial, se mantuvo la presión de poros igual al esfuerzo confinante, se produjo a los 20 m de profundidad una caída brusca en resistencia y rigidez, alcanzando el suelo su estado residual con altas deformaciones.

En la Figura 7 se observa que una vez alcanzada la resistencia residual, el material presenta un comportamiento similar al modelo elasto-plástico, es decir durante los procesos de carga y descarga, la rigidez (pendiente de los lazos de histéresis) tienden a ser constantes, y resistencia cedente igual a la residual. Este hecho sustenta, que en el estado residual, el material aún posee poca resistencia y rigidez, es decir, no se produce una rigidez o resistencia al corte nula, semejante a la que posee el agua y esto permite que se transmitan ondas de corte sísmicas hacia la superficie y el material ofrezca una pequeña resistencia a la deformación.

En las trayectorias de esfuerzos se observa el comportamiento típico de arenas contractivas, es decir la trayectoria de esfuerzos tiende a acostarse hacia la izquierda, no mostrando comportamiento parcialmente contractivo o dilatante. Cabe destacar en este punto algunas definiciones que han sido convenidas para tratar el tema de licuación:

Licuación: denota la condición en que el suelo experimenta grandes deformaciones bajo un estado residual de resistencia al corte, también conocido como falla por flujo, con presión de poros igual al esfuerzo confinante.

Licuación Inicial: denota la condición durante el cual bajo la aplicación de esfuerzos de corte cíclico la presión de poros se iguala con el esfuerzo confinante.

Licuación Parcial: denota la condición en que el suelo después de alcanzar la licuación inicial y durante subsecuentes ciclos de carga y descarga se deforma de manera limitada sin alcanzar el estado residual, también conocido como movilidad cíclica.

Numerosos casos han sido reportados en la literatura de daños a estructuras debido a licuación de suelos, observándose este fenómeno en sitios donde han ocurrido terremotos, con manifestación de volcamiento de edificios, caída de puentes, deslizamiento de presas, agrietamiento de pavimentos, tanques que salen a flote a la superficie, volcanes de arena, entre otros. Ver Fotos 1, 2, 3, 4 y 5.



Foto 1. Volcamiento de edificios en Niigata (Japón) 1964 [17]

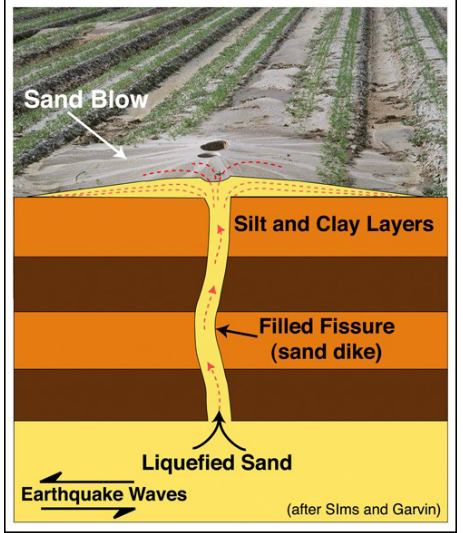


Foto 2. Formación de volcanes de arena por migración de partículas hacia la superficie [17]



Foto 3. Agrietamiento del pavimento por licuación de arenas [17]



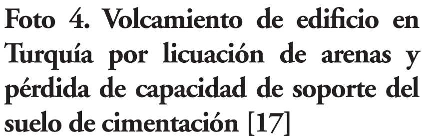

Foto 4. Volcamiento de edificio en Turquía por licuación de arenas y pérdida de capacidad de soporte del suelo de cimentación [17]



Foto 5. Asentamiento diferencial en tanque petrolero producto de licuación de arenas sueltas ante un evento sísmico [17]

4.1.*Evaluación del Potencial de Licuación*

Para determinar si el estrato granular ubicado entre 0 y 7 m de profundidad licuará bajo los escenarios de carga SLE (200 años de periodo de retorno) y DLE (1000 años de periodo de retorno) se recurrió a los resultados obtenidos del estudio particular de sitio obteniendo el valor de CSR con profundidad, el cual resulta de dividir el valor equivalente del esfuerzo cortante cíclico máximo inducido por el sismo (0.65 tmax) entre el esfuerzo vertical efectivo.

La comparación entre los valores de CSR y CRR (cyclic resistance ratio) obtenido de los valores de resistencia de punta del cono medido en tres perforaciones realizadas, aplicando la metodología de Robertson y Cabal [18], para los suelos granulares, permite determinar el Factor de Seguridad contra la licuación definido como:

FS = CRR/CSR
FS: Factor de seguridad contra la licuación
CRR: Resistencia al corte cíclico.
CSR: Razón de esfuerzo cíclico inducido por el sismo.

La representación gráfica de los valores de CSR y CRR, conjuntamente con el Factor de Seguridad (FS) con profundidad para el estrato granular ubicado entre 0 y 7 m de profundidad para los escenarios SLE y DLE se muestran en las Figuras 10 y 11 respectivamente.

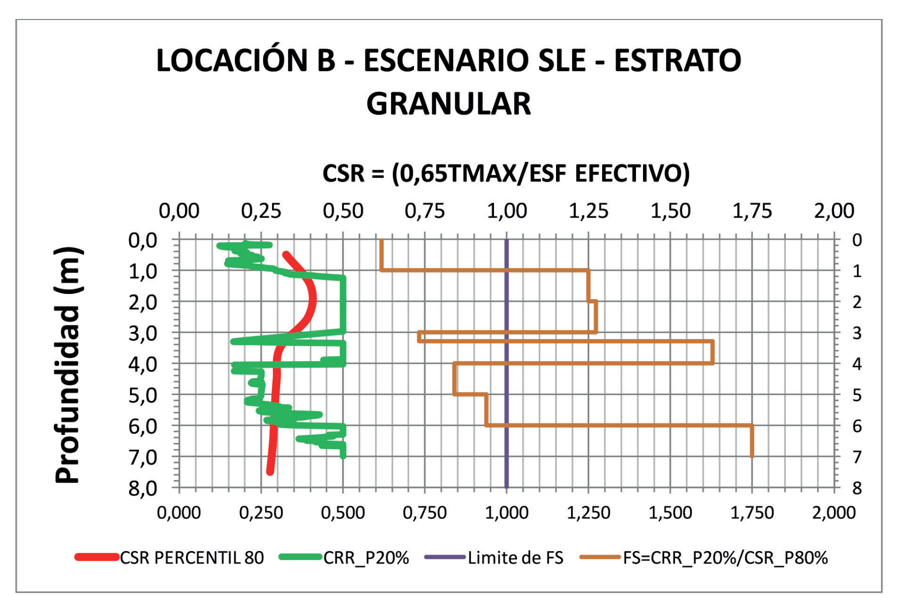


Figura 10. Factor de seguridad contra la licuación del estrato granular – Escenario de carga SLE

Particularmente, bajo el escenario SLE algunos lentes dentro del estrato granular licúan (línea marrón por debajo del Factor de Seguridad de 1 en la Figura 10) y otros no, sin embargo conservadoramente, se va a considerar para efectos de diseño que todo el estrato experimentará licuación.

Bajo el escenario de carga DLE sucede un comportamiento similar al anterior pero abarcando una zona de mayor espesor o altura del estrato licuable como se muestra en la Figura 11.

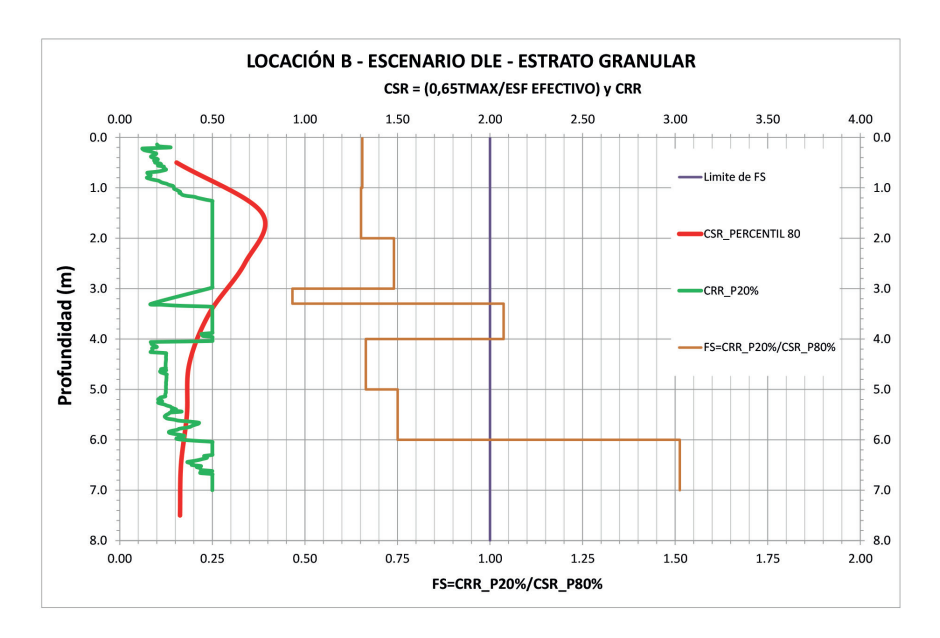


Figura 11. Factor de seguridad contra la licuación del estrato granular – escenario de carga DLE

El estrato granular potencialmente licuable degradará su resistencia al valor de resistencia residual (6 Kpa), que al compararlo con el valor de resistencia pico (23 Kpa) resulta un factor de degradación de resistencia de 0.26. La degradación de rigidez será de 0.1, basado en la medición de pérdida de rigidez, obtenida a partir de numerosos ensayos de corte cíclicos realizados en suelos granulares que han experimentado licuación [19] y [20].

Una vez determinadas los estratos potencialmente licuables y las variaciones de deformaciones equivalentes con profundidad en los demás estratos que componen el perfil litológico, se presentan en las tablas 1 y 2 los factores de degradación de rigidez y resistencia en el campo libre, con los cuales deberán afectarse las curvas p-y o t-z de capacidad portante de pilotes, para tomar en cuenta el efecto de degradación por carga sísmica en los escenarios SLE y DLE respectivamente.

Tabla 1. Factores de degradación de rigidez y resistencia con profundidad – Escenario de carga SLE

Profundidad (m)	Degradación de rigidez	Degradación de resistencia
0 – 7*	0.1	0.25
7 - 15	0.94	1
15 – 25	0.98	1
25 - 100	1	1

* Licuación del estrato granular

Tabla 2. Factores de degradación de rigidez y resistencia con profundidad – Escenario de carga DLE

Profundidad (m)	Degradación de rigidez	Degradación de resistencia
0 – 7*	0.1	0.25
7 - 15	0.88	1
15 – 25	0.90	1
25 – 40	0.94	1
40 – 60	0.98	1
60 - 100	1	1

* Licuación del estrato granular

5. Espectros de Diseño

Haciendo el estudio particular de sitio o de propagación de ondas sísmicas a través de la columna litológica de suelo se han obtenido los espectros de respuesta en superficie. En las Figuras 12 y 13 se muestran los espectros de respuesta para los escenarios de carga SLE y DLE, conjuntamente con el espectro derivado del estudio de amenaza sísmica para períodos de retorno de 200 años y 1000 años respectivamente.

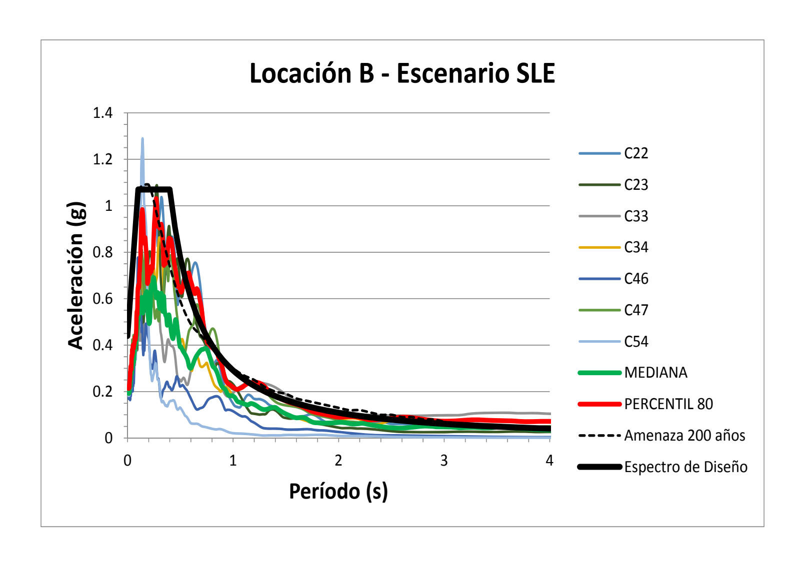


Figura 12. Espectros de respuesta - Escenario de carga SLE

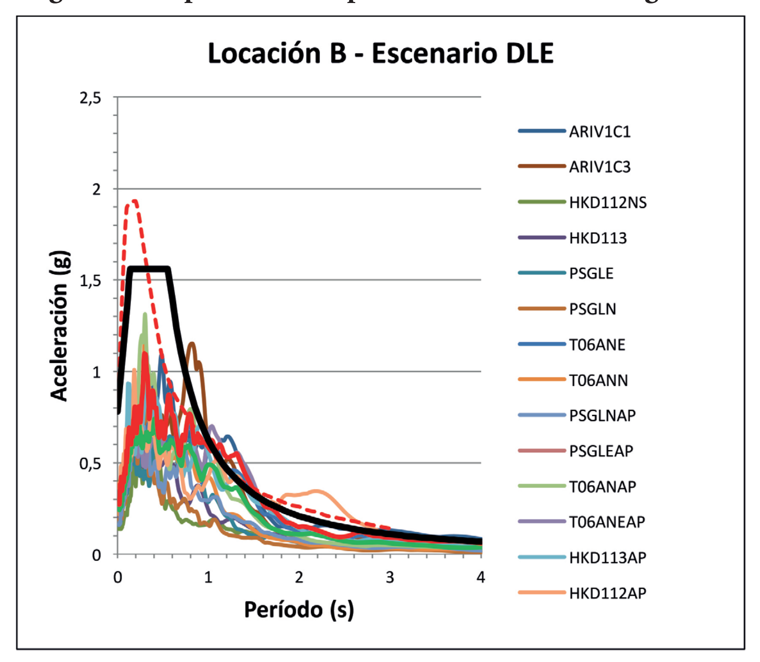


Figura 13. Espectros de respuesta – Escenario de carga DLE

Con los espectros de diseño obtenidos de los estudios particulares de sitio, podrán realizarse análisis dinámicos de las estructuras utilizando el método de superposición modal en estructuras por construir o ya construidas o existentes a fin de cuantificar si se amerita o no la implementación de alguna medida de adecuación o reforzamiento estructural.

6. Conclusiones

En el diseño de estructuras el análisis dinámico es muy importante, por cuanto las cimentaciones así como la superestructura van a estar sometidas a cargas variables en el tiempo, como son los sísmos.

El efecto sísmico se ha tratado considerando los escenarios de carga Strength Level Earthquake (SLE) y Ductility Level Earthquake (DLE), correspondientes a períodos de retorno de 200 años y 1000 años, respectivamente.

Para ello y con base en el estudio de amenaza y los parámetros determinísticos de los posibles sismos que pudiesen ocurrir en la zona de estudio, se hace el análisis de propagación de ondas sísmicas en el subsuelo para ambos escenarios a fin de determinar los valores de esfuerzo y deformación por corte cíclica en campo libre a lo largo de todo el perfil litológico.

Una vez conocidos los valores de esfuerzo y corte cíclicos inducidos por el sismo en el campo libre se determinan los niveles de movilidad cíclica que se producen en los estratos y la susceptibilidad de los suelos granulares a la licuación, y como ello afecta los valores de rigidez y resistencia de los suelos bajo acción dinámica y su influencia en los valores de capacidad por carga axial y lateral de pilotes.

Bajo el análisis particular de sitio, basado en la selección de sismos de la red mundial y el estudio de propagación de ondas en el subsuelo, conjuntamente con el estudio de amenaza, se presenta el espectro de diseño propuesto para períodos de retorno de 200 años y 1000 años en la zona de estudio.

Con los espectros de diseño generados de los estudios particulares de sitio puede evaluarse cuál es el comportamiento esperado de estructuras por construir que se ubican en la cercanía a fallas activas o de aquellas estructuras ya construidas a fin de implementar

las medidas de adecuación o reforzamiento que sean necesarias para que la estructura no colapse aun cuando actúe el sismo más severo.

Referencias

[1] Esmeraldas News. Historia de los terremotos en Ecuador. Disponible en: http://esmeraldasnoticias.blogspot.com/2011/03/historia-de-los-terremotos-en-ecuador.html.

[2] Levantamiento geológico de los efectos del sismo del 16 de abril en la zona entre Esmeraldas y Manta. Disponible en: http://www.igepn.edu.ec/servicios/noticias/1332-levantamiento-geologico-de-los-efectos-del-sismo-del-16-de-abril-en-la-zona-entre-esmeraldas-y-manta.

[3] 16 de abril de 2016: Lecciones aprendidas – Dra. Alexandra Alvarado. Disponible en: http://www.igepn.edu.ec/interactuamos-con-usted/1574-16-de-abril-de-2016-lecciones-aprendidas-dra-alexandra-alvarado.

[4] M. Tamayo, "El Proceso de la Investigación Científica", Editorial Limusa, 3era edición, 1998, pp 37-90.

[5] Base de datos sismológicos de California, Pacific Earthquake Engineering Center (PEER). Disponible en: https://ngawest2.berkeley.edu/

[6] Software Cyclic 1-D. A. Elgamal, Z. Yang, E. Parra and A. Ragheb, 2013. Disponible en: http://soilquake.net/cyclic1d/references.html.

[7] E. Parra, "Numerical Modeling of Liquefaction and Lateral Ground Deformation Including Cyclic Mobility and Dilation Response in Soil Systems", Ph.-D. Thesis, Rensselaer Polytechnic Institute, Troy, New York, 1996.

[8] A. Elgamal, Z. Yang y E. Parra, "Computational Modeling of Cyclic mobility and Post-Liquefaction Site Response", Soil Dynamics and Earthquake Engineering, Vol. 22, pp. 259-271, Elsevier, January, 2002.

[9] R.B. Seed, K. Cetin, R. Moss, A. Kammerer, J. Wu, J.M. Pestana y M. Riemer, "Recent Advances in Soil Liquefaction Engineering and Seismic Site Response Evaluation", Proceedings 4th International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, Symposium in Honor of Prof. W.D.L. Finn, California, March, 2001.

[10] ASCE, "Soil Dynamics and Liquefaction 2000", Proceedings of Sessions of Geo-Denver 2000, Geotechnical Special Publication, No. 107, Denver, Colorado, August, 2000.

[11] Steven L. Kramer (1996). "Geotechnical Earthquake Engineering". University of Washington. Prentice – Hall International Series in Civil Engineering and Engineering Mechanics, pp 184-417.

[12] A. Vásquez y Dobry, "The Behavior of Undrained Contractive Sand and its effect of Earth Structures", Ph. D. Thesis, Rensselaer Polytechnic Institute, Troy, New York, 1988.

[13] W. D. Finn, P. M. Byrne y G. R. Martin "Seismic Response and Liquefaction of Sands", Journal of the Geothecnical Engineering Division, GT-8, 841- 856, August 1976.

[14] W. D. Finn, K.W. Lee y G. R. Martin "An Effective Stress Model for Liquefaction", Journal of the Geothecnical Engineering Division, GT-6, 517-533, June, 1977.
[15] ASTM D-1586, "Standard Test Method for Penetration Test and Split-Barrel Sampling of Soils", Soils and Rocks. American Society for testing and materials.
[16] NEC-SE-DS (2014), "Norma Ecuatoriana de la Construcción. Peligro Sísmico y Diseño Sismo Resistente". Dirección de Comunicación Social, MIDUVI, Ecuador.
[17] UC Davis, Civil and Enviromental Engineering, Earthquake Hazards. Disponible en: https://research.engineering.ucdavis.edu/gpa/earthquake-hazards/.
[18] Robertson, P. K. y Cabal, K. L. (2015). Guide to Cone Penetration Testing for Geotechnical Engineering. (6th). California, EEUU: Gregg Drilling & Testing, Inc.
[19] Stark, T. and Mesri, G. (1994) "Undrained Shear Strength of Liquefied Sands for Stability Analysis". Journal of Geotechnical Engineering. Vol. 118, No. 11, 120(7), 1287–1287.
[20] K. Ishihara (1993), "Liquefaction and flow failure during earthquakes". Géotechnique, Volume 43, Issue 3, pages 351 –451.

Sobre el autor



José Domingo Alviar Malabet (Y´1962-M´06). El Ing. José Alviar se ha desarrollado dentro de la Ingeniería Civil especializándose en el área de la geotecnia y estructuras, así como en la docencia por más de 25 años. Ha gerenciado, liderizado y participado en la supervisión de proyectos geotécnicos y estructurales costafuera y en tierra firme, para el sector público, privado y para la industria petrolera. En el año 1985 se gradúa de Ingeniero Civil en la Universidad Católica Andrés Bello, en el año 1993 recibe el Magister en Ingeniería Civil (Geotecnia) Universidad Simón Bolívar y en el año 2014 obtiene su grado de Ph.D en Estructuras de la Universidad Central de Venezuela. Actualmente es profesor de la Universidad Politécnica Salesiana, Sede Quito, campus Sur en la carrera de Ingeniería Civil. Su principal línea de investigación está relacionada con los problemas de interacción suelo-estructura, el diseño sismo resistente de cimentaciones superficiales y profundas, los estudios de propagación de ondas sísmicas en el subsuelo y la generación de espectros de diseño.

CAPÍTULO V

Diseño de una arquitectura heterogénea para la gestión eficiente de recursos FPGA en un cloud privado

Desing of a heterogeneous architecture for the efficient management of FPGAs in a Cloud

Julio Proaño Orellana*, María Blanca Caminero†, Carmen Carrión§
*Universidad Politécnica Salesiana, †Universidad de Castilla La Mancha
*email: jproanoo@ups.edu.ec
†email: mariablanca.caminero@uclm.es
§email: carmen.carrion@uclm.es

Resumen

El uso eficiente de recursos en entornos de computación en la Nube es uno entre los muchos temas que han atraído a muchos científicos en los últimos años. Hoy en día existe una amplia variedad de recursos informáticos disponibles en las infraestructuras de Cloud Computing, como CPUs multi-core, GPUs y FPGAs. En particular, arquitecturas basadas en aceleradores que usan FPGAs pueden proporcionar un importante valor agregado de negocio en entornos Cloud. Los dispositivos FPGA ofrecen un mejor desempeño debido a su gran capacidad de procesamiento con latencia predecible y bajo el consumo de energía. Sin embargo, la gestión de estos recursos en un entorno Cloud no es trivial, y es el foco de este trabajo. En este trabajo, se presenta una arquitectura destinada a integrar FPGAs como aceleradores en un Cloud privado. La propuesta toma en cuenta requisitos de calidad de servicio (QoS). Así, los indicadores de QoS están relacionados con los requerimientos de los usuarios. La arquitectura propuesta busca la combinación adecuada entre CPUs y FPGAs en un entorno Cloud.

Palabras clave

Cloud Computing, energía, FPGA

Abstract

The efficient use of resources in cloud computing environments is one of the many topics that have attracted many scientists in recent years. Today there is a wide variety of computing resources available in Cloud Computing infrastructures, such as multi-core CPUs, GPUs and FPGAs. In particular, accelerator-based architectures using FPGAs can provide

significant business value in Cloud environments. FPGA devices offer better performance due to their high processing capacity with predictable latency and low power consumption. However, the management of these resources in a Cloud environment is not trivial. In this work, an architecture designed to integrate FPGAs as accelerators in a private Cloud is presented. The proposal takes into account quality of service (QoS) requirements. Thus, the QoS indicators are related to the requirements of the users. The proposed architecture seeks the right combination between CPUs and FPGAs in a Cloud environment.

Keywords

Cloud Computing, energy, FPGAs

1. Introducción

En la actualidad el paradigma de Cloud Computing es un tema de gran interés tanto para la academia como para la industria. La industria de Cloud representada por los proveedores de servicios invierte grandes cantidades de dinero en nuevas infraestructuras con el objetivo de satisfacer la creciente demanda de recursos computacionales. Sin embargo, problemas como por ejemplo: la gestión eficiente de recursos, la seguridad, la confidencialidad de la información y el excesivo consumo de energía son cuestiones que aún no se han resuelto del todo. Así, la academia es la que se enfoca en resolver estos problemas a través de la investigación científica. [1]

Los entornos Cloud están conformados por dos elementos fundamentales, proveedores y clientes. Los proveedores, siendo los propietarios de los recursos, buscan maximizar sus ganancias ofreciendo recursos computacionales como un servicio a bajo coste. Por otro lado, los clientes buscan ejecutar sus aplicaciones sin tener que incurrir en gastos de manteniendo de infraestructura; dicho en otras palabras ellos únicamente buscan pagar por el uso de los recursos.

La interacción entre estos dos actores sigue un protocolo simple. Los clientes envían peticiones a los proveedores para utilizar recursos computacionales, los proveedores por su parte luego de procesar dichas peticiones seleccionan y encapsulan los recursos solicitados en una máquina virtual. Finalmente, el cliente recibe el acceso a esta máquina como respuesta a su requerimiento. [2]

Los proveedores de servicios Cloud son los responsables de asegurar una calidad de servicio (QoS) en la cual se consideran paráme-

tros como: la disponibilidad del servicio, el rendimiento, la latencia, etc. [3] Con el objetivo de asegurar a los clientes que estos requerimientos de calidad se cumplan, clientes y proveedores negocian un acuerdo denominado en inglés Service Level Agreement (SLA). [4]

Clientes y proveedores persiguen diferentes objetivos, por un lado, los clientes buscan utilizar recursos computacionales para sus aplicaciones con el mínimo costo, mientras los proveedores buscan maximizar sus ingresos manteniendo una calidad de servicio adecuada y disminuyendo sus gastos. Dicho de otra manera, los proveedores buscan ser eficientes en la administración de recursos y el uso de la energía. Sin embargo, según estudios publicados por Internet World Stats las regiones como África, Latinoamérica y el Caribe han mostrado un crecimiento del 7415% y 2029% en los años comprendidos entre 2000 y 2016. Como consecuencia, también el coste de consumo de energía y disipación de calor se han incrementado de manera exponencial. Un ejemplo claro se muestra en [5] donde los autores sostienen que el 42% del presupuesto total de los servidores de Amazon EC2 se corresponde con la energía tanto en generación como en disipación. De hecho, de acuerdo a un estudio realizado por McKinsey & Company [6] en promedio el 80% de la energía utilizada por los grandes centros de datos es únicamente para mantener a estos en estado de reposo y el 20% restante es usado para operaciones de cómputo.

Según un informe realizado por el departamento de Energía Lawrence Berkley National Laboratory [7] en el 2016, una estrategia como el uso de hardware más eficiente podría disminuir el consumo de energía en un 45% para el 2020. Este reporte menciona que el uso de infraestructura más eficiente en almacenamiento, red y procesamiento el fundamental para disminuir el excesivo consumo de energía de los grandes centros de datos.

Desde la industria, los principales proveedores de Cloud como Google, Amazon han optado como estrategia la utilización de infraestructuras más eficientes como el uso de dispositivos FPGA [8]. Como resultado han logrado reducir el consumo de energía. Además, han trasladado sus infraestructuras a lugares con bajas temperaturas, para disminuir el consumo de energía que se utiliza para enfriar sus granjas de servidores.

Como se explicó anteriormente es muy importante establecer una arquitectura que permita por un lado satisfacer la creciente demanda de recursos pero al mismo tiempo contralar el consumo de energía.

Este trabajo propone una arquitectura que aborda la integración de hardware eficiente como FPGAs dentro de un paradigma de Cloud Computing. La propuesta se enfoca en la gestión de recursos heterogéneos que incluyen el uso de FPGAs como aceleradores de código en la ejecución de tareas en un entorno de Cloud privado.

Las principales contribuciones de este artículo son: una revisión de la literatura enfocada en la integración y gestión de FPGAs dentro de entornos Cloud y el desarrollo teórico de una arquitectura que tenga como objetivo la integración de FPGAs en un Cloud privado. Finalmente, una evaluación como prueba de concepto de la propuesta.

El artículo se encuentra organizado de la siguiente manera, en la Sección 2 se muestra una revisión de los conceptos teóricos que soportan la propuesta. En la Sección 3, se exponen los principales trabajos relacionados con nuestra propuesta. En la sección 4, se dan detalles sobre la implementación y funcionamiento de la propuesta. En la sección 5 se muestra una evaluación como prueba de concepto de la arquitectura. Finalmente, en el capítulo 6 se discuten algunas conclusiones de este trabajo.

2. Marco teórico

A. *Arquitectura Cloud*

Desde el punto de vista de la computación distribuida, el paradigma de Cloud sigue una arquitectura por capas. Las tres principales capas son: la capa de plataforma de gestión, la capa de virtuzalización y la capa de hardware como se puede observar en la Figura 1. [9]

La capa de hardware está compuesta por un grupo de recursos tales como: servidores, dispositivos de red, fuentes de energía, sistemas de ventilación y disipación de calor. Además, cabe destacar que cada servidor se encuentra compuesto, por lo menos, por procesadores, memoria, discos de almacenamiento y tarjetas de red. También se pueden encontrar dispositivos especiales como por ejemplo co-procesadores, procesadores gráficos (GPUs), FPGAs, etc.

La capa de virtualización usa máquinas virtuales para partir los recursos computacionales disponibles y así facilitar su compartición y gestión. La virtualización permite crear un entorno virtual de manera segura y asilada con el objetivo de ejecutar una aplicación o un grupo de aplicaciones determinadas.

Desde el punto de vista de Cloud, los proveedores de servicio necesitan encontrar mecanismos para realizar una gestión eficiente y segura de los recursos computacionales. Y además, asegurar que dicho mecanismo sea capaz de escalar.

Los proveedores utilizan la virtualización para entregar servicios y/o desplegar servicios a los clientes de manera rápida y segura. A pesar de que el término de virtualización sea comúnmente relacionado con la emulación de un hardware, esta tecnología es muy útil para la compartición y gestión de recursos en un centro de datos. El mecanismo de gestión, a nivel macro se lo realiza a través de un software conocido como Gestor de Virtualización en inglés Virtual Machine Manager VMM. Existen diferentes tipos de VMMs, los más usados en entornos Unix son: KVM (Kernel-based Virtual Machine) y XEN, mientras que para los sistemas que se ejecutan en Windows se tiene Hyper-V y VMware vSphere. [10]

A pesar de las ventajas para la gestión, aislamiento y seguridad que ofrece la virtualización, esta tecnología presenta un sobrecoste en el rendimiento del sistema debido a que el VMM debe multiplexar los procesos del sistema operativo local y los procesos del sistema operativo simulado. Sin embargo, existen técnicas como por ejemplo la paravirtualización que ayuda a reducir este sobrecoste de rendimiento.

Finalmente, grandes empresas tecnológicas como Intel y AMD han desarrollado extensiones para mejorar el rendimiento de entornos virtuales. En particular, Intel VT-x [11] y AMD SVM [12]. Estos mecanismos permiten acelerar los entornos virtuales a través de funciones especiales y la gestión de memoria.

La última capa denominada plataforma opera a nivel de nodo de un servidor y se encarga de gestionar las distintas máquinas virtuales desplegadas en diferentes nodos de un centro de datos.

Esta capa está compuesta por un software que tiene la capacidad de conocer todos los recursos disponibles de varios nodos en un entorno Cloud. Además, esta capa es capaz de desplegar, moni-

torizar y gestionar una gran cantidad de máquinas virtuales desplegadas en la capa de virtualización. Dicho de otro modo, esta capa es la responsable de dirigir de manera efectiva la creación, monitorización, gestión, despliegue de toda la infraestructura virtual en un entorno de Cloud.

En la actualidad tanto en la academia como en la industria los VMMs más usados en entornos de Cloud son: Open Stack [13], Eucalyptus, VMware y Open Nébula [14]. Cada uno de estos contiene una seria de características que los investigadores aprovechan en base a sus necesidades.

En el siguiente apartado se detallan, aspectos sobre el paradigma de Cloud Computing basado en tecnologías de virtualización.

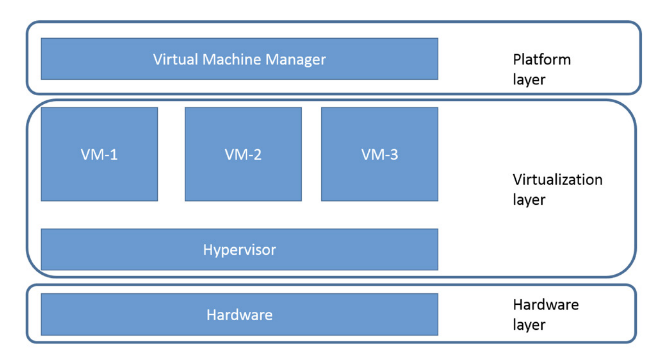


Figura 1. Arquitectura de Cloud Computing

B. Definición y modelos de despliegue en Cloud

De acuerdo con el Instituto Nacional de Estándares y Tecnología en inglés Nation Institute of Standars and Technology (NIST) [15], Cloud Computing es "un modelo ubicuo bajo demanda que comparte recursos computacionales (red, servidores, almacenamiento, aplicaciones y servicios) que pueden ser rápidamente provistos y liberados con una gestión mínima por parte de los proveedores."

Según la definición anterior Cloud Computing presenta una serie de caracterizas tales como:

- **Bajo-demanda**.- los proveedores de servicios Cloud permiten auto-aprovisionamiento de recursos de manera autónoma.

- **Acceso ubicuo**.- el acceso a los servicios de Cloud es a través de Internet.
- **Multi-usuario**.- los recursos en Cloud son compartidos por un sin número de usuarios de manera transparente y eficiente.
- **Elástico y escalable**.- Cloud tiene la capacidad de adicionar y liberar recursos de manera dinámica de acuerdo a los requerimientos de los clientes.
- **Monitorización de uso de recursos**.- para controlar, optimizar y con fines de facturación, las arquitecturas Cloud ofrecen herramientas para la monitorización del uso de recursos.

Por otro lado, el paradigma de Cloud ofrece tres tipos modelo de entrega de servicios que son: Software-como-un-Servicio (SaaS), Plataforma-como-un-Servicio (PaaS) e Infraestructura-como-un-Servicio (IaaS).

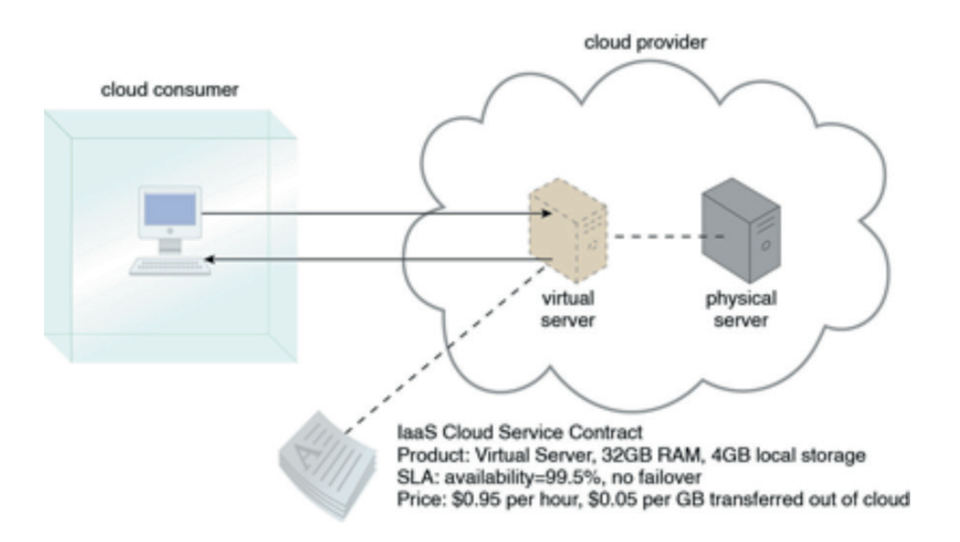


Figura 2. Infraestructura como servicio (IaaS) [3]

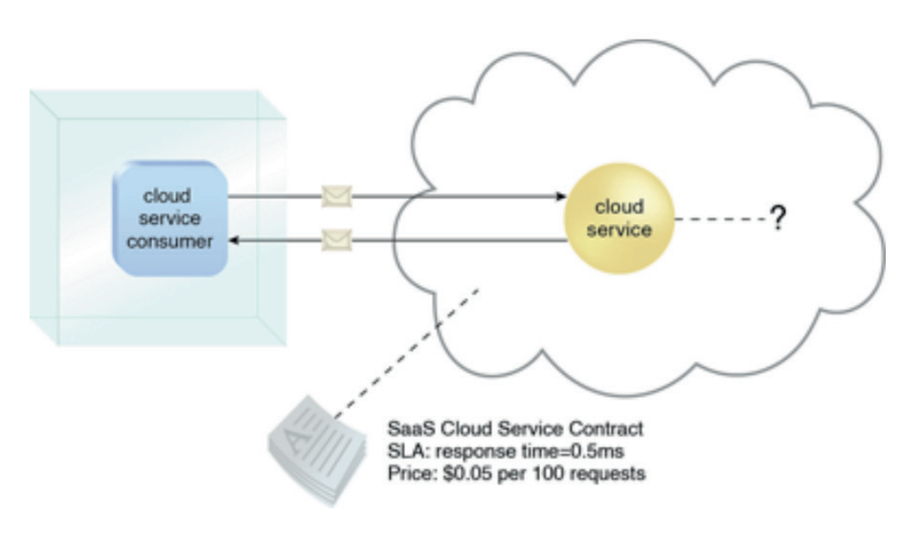


Figura 3 Software como servicio (SaaS) [3]

ready-made environment

cloud consumer

?

virtual server

cloud provider

PaaS Cloud Service Contract
Product: application server + DMBS platforms
SLA: availability=99.5%, auto-scaling
Price: $0.45 per hour (500,000 requests)

Figura 4. Plataforma como servicio (PaaS) [3]

En el modelo de *IaaS* como se muestra en la Figura 2, los proveedores ofrecen a sus clientes recursos computacionales (procesamiento, servicios de red, almacenamiento) encapsulados en una máquina virtual, por ejemplo: Amazon Web Service, Microsoft Azure, Google Cloud, Rackspace, etc.

En el caso de *SaaS* (ver Figura 3), los proveedores ofrecen a sus clientes una lista de aplicaciones que están en el Cloud, algunos ejemplos son: Twitter, Facebook, Flickr, Google Docs.

Finalmente en la Figura 4, se muestra el modelo PaaS. En este modelo, se ofrecen plataformas para el desarrollo de aplicaciones, como por ejemplo: Amazon Web Service (AWS), Windows Azure, Google App Engine.

Finalmente, todos estos servicios de Cloud mencionados anteriormente pueden ser desplegados en entornos *Privados, Públicos, Comunitarios e Híbridos*. A continuación se da un breve resumen de los distintos modelos de despliegue de Cloud.

La Figura 5 muestra el despliegue en un entorno privado, cuya principal característica es que los recursos solamente pueden ser accedidos por los miembros de la misma organización.

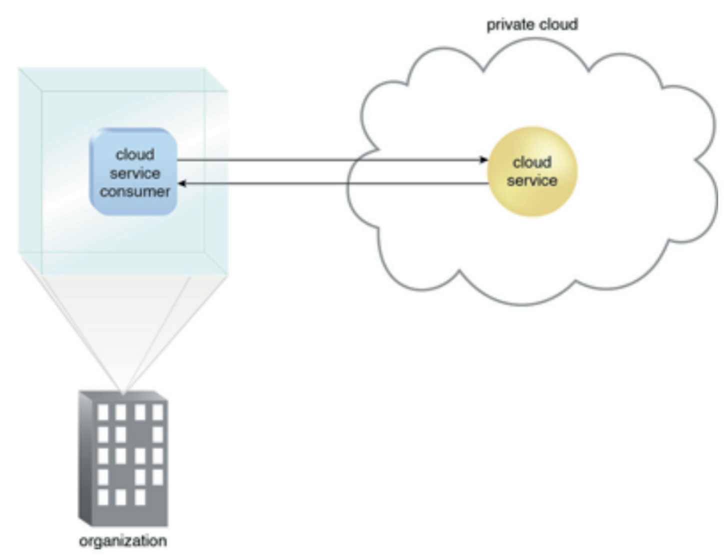


Figura 5. Ejemplo de un Cloud Privado [3]

Por otro lado, la Figura 6 muestra la arquitectura para un modelo de Cloud *Público* en el cual los recursos computacionales se ofrecen al público en general a través de Internet.

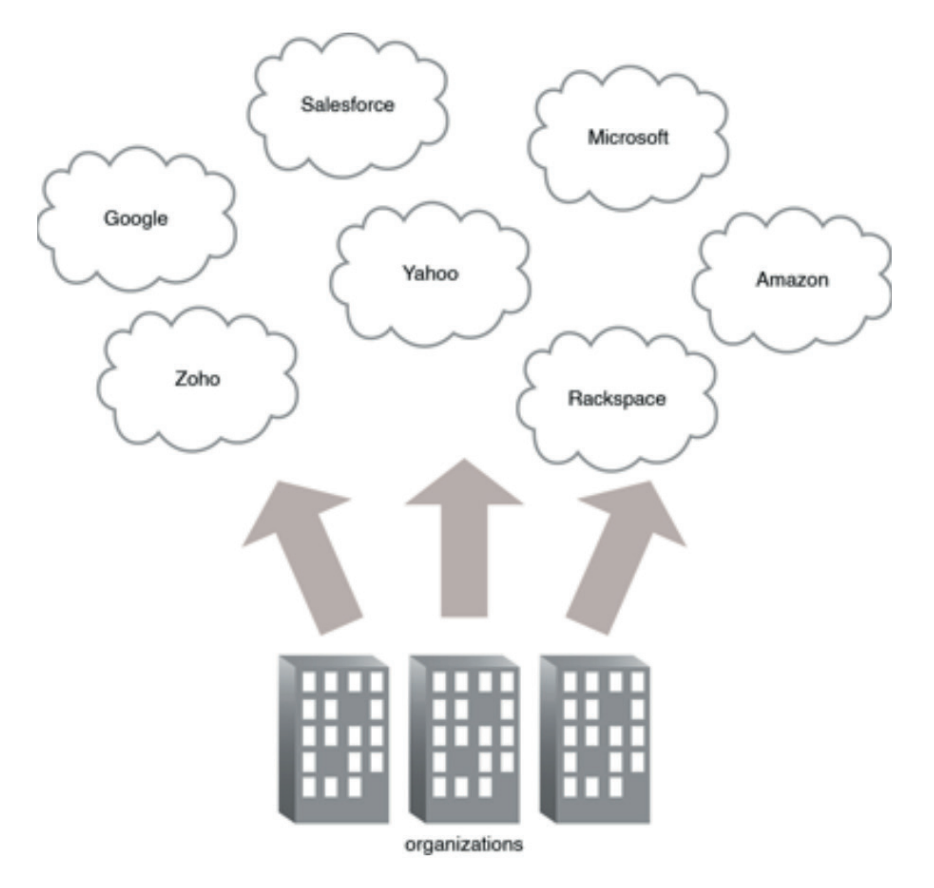


Figura 6. Ejemplo de Cloud Público [3]

Finalmente, la Figura 7 muestra un modelo que combina dos o más de los modelos mencionados anteriormente.

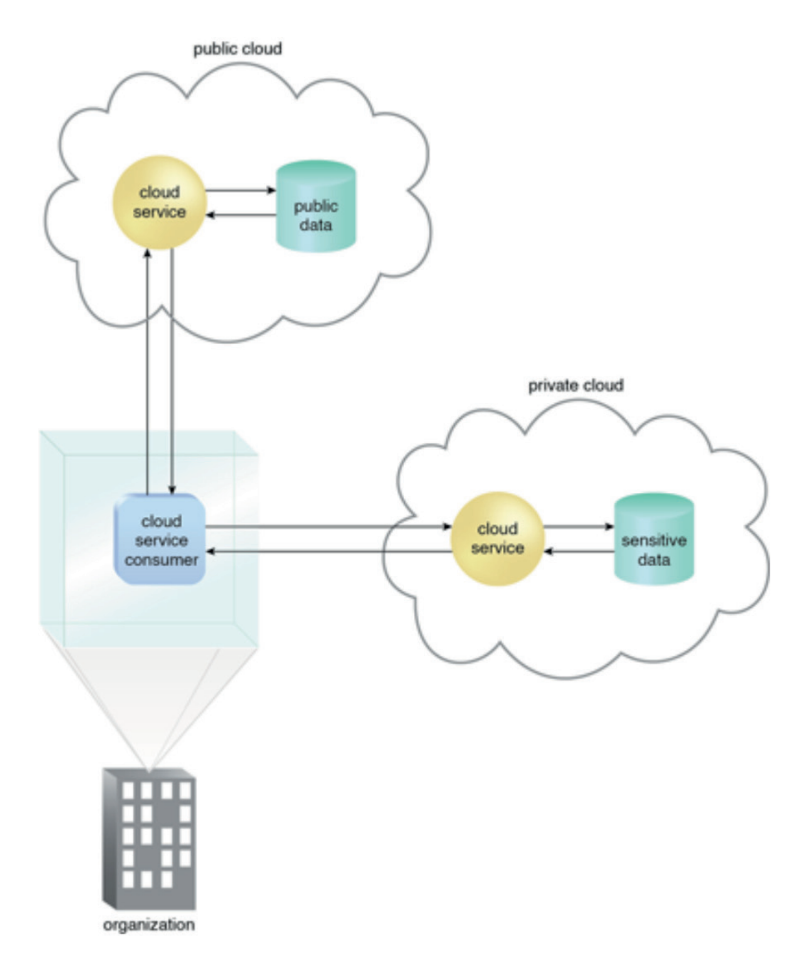


Figura 7. Ejemplo de un Cloud Híbrido [3]

C. Calidad de Servicio (QoS) y asignación de recursos en Cloud

En la sección anterior se mencionaron la definición de un Cloud y los modelos servicio y despliegue. En esta sección nos referiremos a conceptos fundamentales de calidad de servicio y las diferentes técnicas usadas para la gestión de recursos en Cloud.

Los entornos Cloud ofrecen a sus clientes una gran cantidad de recursos disponibles bajo demanda siguiendo un modelo pay-as-you-go, pago por lo que uso [16]. Desde el lado del proveedor, los

servicios tienen cumplir con una cierta calidad. Estos parámetros de calidad tales como: rendimiento, latencia, disponibilidad son incluidos dentro del Acuerdo de Nivel de Servicio (SLA) [17].

La calidad del servicio en Cloud está estrechamente relacionada a la gestión de recursos del sistema. Una gestión eficiente permite a los proveedores aprovechar al máximo sus infraestructuras. Sin embargo, debido a la complejidad que presenta un Cloud, esta tarea no es trivial. Uno de los problemas más abordados es la asignación eficiente de recursos.

La asignación de recursos es una tarea que debe ser realizada de manera autónoma y bajo demanda. Por lo general se definen distintas políticas para llevar esta tarea. Las políticas más usadas para la asignación de recursos en Cloud incluyen, el control de admisión, la eficiencia energética, los modelos estadísticos, el balanceo de carga, criterios de SLA y el uso de dispositivos dedicados como FPGAs. [18]

La política de control de admisión es preventiva y su objetivo es evitar la sobrecarga de recursos y violaciones del SLA. Esta política se basa en la relación entre el conocimiento previo acera del comportamiento del sistema y la carga [19]. Esta política puede resultar muy compleja para entornos muy dinámicos y puede causar una sub-utilización del sistema. Sin embargo, trabajos [20], demuestran que la combinación de estas políticas con criterios sobre aprovisionamiento (Over-booking) suelen dar un mejor rendimiento.

Por otro lado, las políticas basadas en eficiencia energética y modelos estadísticos buscan un balance entre la asignación de recursos, mantener una tasa aceptable de SLAs cumplidos y disminuir el consumo de energía. Proyectos como TANGO [21], se enfocan en la implementación de algoritmos de eficiencia energética orientado a computación de alto rendimiento. Trabajos como [5], [22], [23], usan algoritmos que consideran el consumo de energía.

Las políticas que consideran el balanceo de carga tienen como objetivo distribuir la carga a través de los recursos disponibles del sistema. Estas políticas integran criterios de migración de tareas a diferentes recursos dependiendo de la carga. Por ejemplo, si el sistema detecta que un nodo se encuentra subutilizado y otro sobreutilizado, el sistema se encarga de migrar las tareas desde el nodo sobreutilizado al nodo subutilizado, de tal manera que se consiga un

equilibrio. Sin embargo, la migración de tareas podría provocar una saturación de la red. [18]

Las políticas que toman en cuenta los SLAs [19] se enfocan en mapear el requerimiento de cada uno de los clientes y los recursos disponibles que pueden satisfacer el SLA. Por lo general se basan en algoritmos que realizan predicciones y toma de decisiones relacionados con la Inteligencia Artificial.

Finalmente, las políticas relacionadas con la utilización de dispositivos más eficientes son nuevas en el ámbito de Cloud. En este tipo de mecanismo las peticiones de los clientes son clasificadas en diferentes categorías dependiendo de los requerimientos para después ser asignadas a los recursos que mejor se adapten. Además, estos recursos poseen características muy favorables con respecto al uso de la energía y rendimiento. Sin embargo, es una tarea muy compleja encontrar los mejores recursos para todas las tareas por lo cual se puede tener algunos recursos subutilizados.

D. *Técnicas de Planificación de recursos en Cloud*

Desde el punto de vista de la industria las técnicas de planificación de recursos son utilizadas para maximizar los ingresos de los proveedores y reducir el tiempo de retorno de la inversión (ROI). El mecanismo más utilizado para conseguir este objetivo es compartir de manera eficiente los recursos computacionales disponibles. Sin embargo, esta tarea no es trivial debido a la gran variabilidad que presentan los entornos de Cloud. Según [24], la clave para conseguir el objetivo antes mencionado está en diseñar un sistema de planificación que sea eficiente, lo suficientemente justo e inteligente con la mínima interacción humana.

De manera general algunos mecanismos para la gestión se basan en sistemas de colas con prioridades para determinar ¿Qué recurso? va a ¿Dónde? y ¿Cuándo? Esta estrategia básicamente elabora listas dinámicas y ordena sus elementos según un criterio de prioridad.

Las técnicas de planificación de recursos se pueden implementar siguiendo una arquitectura por capas. Así, cada capa es responsable únicamente de una función específica. Por ejemplo: en un esquema de planificación con tres capas: Aplicación, Virtualización y

Despliegue. La primera capa es la responsable de clasificar las peticiones de los clientes de acuerdo a una determinada política como QoS. La segunda capa es responsable, por su parte, por su parte de gestionar los recursos virtualizados haciendo un mapeo entre los recursos físicos y los virtuales. Finalmente, la última capa es responsable de repartir las tareas a diferentes nodos del Cloud.

La planificación de recursos puede ser vista como un problema de optimización multiobjetivo. Por un lado minimizar el coste y maximizar la utilización de los recursos en el tiempo.

En [25], los autores clasifican las estrategias de planificación de recursos en aquellas que consideran el rendimiento (performance) y en las que consideran el coste. Las estrategias que consideran el rendimiento se dividen en First-come First-serve (primero en llegar es el primero en ser atendido), balanceadores de carga. Por otro las estrategias que toman en cuenta el coste se dividen en máximo rendimiento, mínimo coste de operación y la combinación de las dos.

A continuación se hace una revisión de trabajos relacionados con la integración, gestión y planificación de recursos como FPGAs dentro de entornos Cloud.

3. Trabajos relacionados

A. Recursos heterogéneos en Cloud

Existen diferentes tipos de clientes que usan los recursos de Cloud como por ejemplo, científicos, desarrolladores, investigadores etc. Todos ellos, necesitan un conjunto de recursos específicos creando una demanda muy variada. Esta diversidad de clientes demanda también una amplia gama de servicios también muy diversa. Así, los proveedores deben adaptarse a través de la adquisición de más y diversa infraestructura. Sin embargo, el hardware está limitado a factores como la densidad de potencia [W/mm^2] y la complejidad. La diversidad obliga a los proveedores en integrar arquitecturas heterogéneas. La integración de recursos heterogéneos puede ayudar a mejorar la eficiencia en el uso de los recursos a través de un mapeo de las necesidades de los clientes con los recursos más adecuados que contienen caracterizas específicas muy favorables para satisfacer un

SLA. Estas características pueden ser, por ejemplo, un mejor desempeño en el procesamiento gráfico mediante el uso de GPUs, procesadores usados como aceleradores de código incluso hardware dedicado para ciertas tareas específicas mediante el uso de FPGAs.

B. FPGAs en Cloud

Los dispositivos de hardware FPGA usados como coprocesadores ofrecen una mejora significativa para muchas aplicaciones [26]. Así, existen algunos esfuerzos para integrar FPGAs dentro de ámbito tradicional de las aplicaciones informáticas. Algunas soluciones han sido proporcionadas por la industria, tal como PicoComputing, Convey y Xillybus [27]. Estos productos conectan aplicaciones informáticas con la FPGA a través de una interfaz propietaria desarrollada bajo sus protocolos propios. Mitrion-C, Open Bio Project [28] intentan acelerar aplicaciones bioinformáticas portando secciones críticas de sus aplicaciones hacia FPGAs.

Por otro lado, arquitecturas de código abierto como RIFFA [29] proporcionan una plataforma muy práctica para usar dispositivos FPGA de Xilinx como aceleradores de código. Sin embargo, estos ejemplos no disponen de una arquitectura para desarrollar aplicaciones paralelas y distribuidas que usen múltiples nodos interconectados entre sí.

En otros trabajos los dispositivos FPGAs son utilizados solamente como elementos computacionales formando un clúster[30]. Sin embargo, no todas las aplicaciones pueden ser aceleradas de forma efectiva usando FPGAs. Por ejemplo, para aplicaciones basadas en operaciones con números de coma flotante o aplicaciones que requieren una frecuencia de reloj elevada, los procesadores gráficos GPUs son más eficaces como aceleradores de código.

El uso de GPUs como recursos computacionales dentro de infraestructuras de Cloud ha surgido en los últimos años. Por lo tanto, hoy en día algunas compañías ofrecen los servicios de GPU a sus clientes. Por ejemplo, Amazon EC2 [8] soporta el uso de GPUs para flujos de trabajo diseñados bajo CUDA y OpenCL [31].

El uso de FPGAs en sistemas Cloud cada vez va tenido mucho éxito. Así, empresas como Nimbix [32] ofrecen un sistema de pago

en la nube con una variedad de plataformas para aceleración de código. Recientemente, esta compañía ha lanzado JARVICE, que es una plataforma que incluye la disponibilidad GPUs, DPS y FPGAs para un Catálogo de aplicaciones y además una API o un acceso para enviar tareas a través de una línea de comandos.

Por otra parte, hay algunos trabajos que combinan aceleradores (FPGAs, GPUs) y CPUs dentro de nodos en un clúster como AXEL, AXEL combina nodos heterogéneos con el modelo de programación MapReduce obteniendo como resultado una plataforma de computación de alto rendimiento y de bajo coste. En este trabajo el autor destaca que el mayor inconveniente es la dificultad de implementación de los diseños dentro de las FPGAs en comparación con la programación que ofrece por ejemplo CUDA para GPUs. Además, investigaciones recientes destacan el uso de la reconfiguración dinámica de los nuevos modelos de FPGAs. Por ejemplo en [33] se presenta un servicio de reconfiguración. Este servicio está basado en la administración eficiente de la funcionalidad de la reconfiguración parcial de dispositivos FPGAs, obteniendo así una gran disminución en el tiempo de configuración del dispositivo y optimizando al máximo el uso de recursos.

La investigación que más se acerca a nuestra visión es el proyecto Europeo FP7 HARNESS (Hardware and Network Enhanced Software Systems for Cloud Computing) Proyecto Europeo FP7. El objetivo de este proyecto es desarrollar una plataforma mejorada con acceso a una variedad de recursos computacionales, de comunicación y de almacenamiento. La plataforma consiste en una serie de componentes que pueden tener múltiples implementaciones. A continuación, una aplicación puede ser desplegada de diferentes maneras sobre los recursos obteniendo costos diferentes, mayor rendimiento y características de uso. Sin embargo, en este trabajo se enfoca como una extensión de los trabajos antes mencionados en el cual se toma en cuenta la eficiencia energética.

C. FPGAs en IaaS

Existen algunos trabajos en los cuales se integran FPGAs dentro de entornos Cloud para ofrecer estos como un servicio. En [34]

los autores proponen una arquitectura para proveer FPGAs como servicio en el Cloud de IBM. Esta propuesta está implementada sobre el gestor de Cloud OpenStack [13] y su característica principal es el uso de la reconfiguración parcial. La reconfiguración parcial es una característica de las FPGAs modernas y permite compartir dividir el hardware para ser utilizado por varios usuarios de manera simultánea. Existen trabajos similares a este en los cuales se ofrece FPGAs como parte de una arquitectura genérica a través de una extensión del gestor OpenStack [13].

En [35] se propone un hypervisor para la integración de FPGAs virtuales. En este trabajo el hypervisor tiene acceso a una base de datos que contiene la información de los dispositivos FPGA y este se encarga de su gestión e base a la demanda de los clientes.

Todos estos trabajos [36], [37], [38] se enfocan en ofrecer recursos FPGA como un servicio.

Desde el punto de vista de ofrecer software como un servicio (SaaS), el uso de FPGAs como aceleradores de código a tomando importancia recientemente. En 2014, Microsoft lanzó su proyecto Catapult [39], el cual se enfocó en acelerar con FPGAs el algoritmo Page-rank usado para búsquedas en Bing. Los resultados obtenidos fueron de alrededor de 95% en el rendimiento usando únicamente 22.7 [W] adicionales.

En otro trabajo los autores modificaron la arquitectura de Apache Hadoop [40] para incluir el uso de FPGAs como aceleradores de código. En este trabajo se utilizaron las FPGAs en conjunto con nodos workers. En un trabajo similar [41], se utilizaron FPGAs para acelerar los algoritmos tanto de los nodos mappers como reducers. En este trabajo la estrategia de gestión de recursos se realizó mediante el uso de un sistema multicolas para la asignación de tareas.

En [41], se utilizan FPGAs para acelerar algoritmos de Machine Learning como: K-means, Naive Bayes, k-Nearest Neighbours, Support Vector Machine (SVM). Los resultados indican que en promedio la arquitectura con FPGAs puede alcanzar una aceleración de 2.27x y con un consumo menor de 3.7x.

Existen también productos comerciales enfocados a la aceleración de aplicaciones en Cloud como por ejemplo Ryft3.

Este producto puede alcanzar una aceleración de hasta 100x y una reducción de energía total de 70%.

A continuación se expone la arquitectura propuesta para una integración eficiente de FPGAs dentro de un entorno de Cloud Computing privado.

4. Arquitectura propuesta

En esta sección se describe una arquitectura heterogénea para la construcción de servicios de infraestructura bajo demanda en los sistemas de Cloud Computing basada en [37], [36]. La arquitectura intenta abordar los desafíos de despliegue de una infraestructura como servicio (IaaS) en un centro de datos privado, en el que los dispositivos FPGA son considerados como recursos computacionales virtuales.

Las FPGAs virtuales son elementos computacionales administrados por el sistema junto con las máquinas virtuales que se ejecutan en los distintos procesadores (CPUs). Para ser más precisos, los dispositivos FPGA son usados como aceleradores en la ejecución de aplicaciones. En el resto de la sección se describe las principales características de la arquitectura que proponemos. Antes de explicar la propuesta, es muy importante mencionar que cada nodo se compone de al menos una CPU, y que pueden incluirse algunos elementos COTS extra. La CPU es la encargada de controlar el acceso y la asignación de tareas a los aceleradores. Por lo tanto, hace falta una conexión física entre la máquina virtual y los elementos de aceleración de hardware. Se propone usar el bus de datos PCI exprés ya que este es el que proporciona el mejor rendimiento en cuanto a la cantidad de datos que se pueden enviar/recibir por unidad de tiempo. Además cabe destacar que para mejorar aún más el intercambio de datos entre máquinas virtuales y hardware de aceleración pretendemos usar la tecnología de virtualización que proporciona Intel VT-d [11]. Esta tecnología permite que los dispositivos de hardware accedan de manera directa a las máquinas virtuales, reduciendo el tiempo de inactividad y aumentando el rendimiento productivo por una mejor utilización de los recursos del centro de datos.

La mayoría de los entornos Cloud utilizan un software de administración de máquinas virtuales conocido como Virtual Machi-

ne Manager (VMM). Este elemento es el encargado de preparar, desplegar y monitorizar todos los recursos virtuales en un entorno Cloud. La arquitectura propuesta está desarrollada sobre el gestor VMM y permite utilizar las funcionalidades de este elemento y además extender dichas funcionalidades a la explotación de elementos de hardware adicionales como FPGAs o GPUs. Para ello se ha diseñado dos módulos principales, que son: Administrador de Hardware de Aceleración (HAM, por sus siglas en ingles) y una base de datos que hemos denominado Catálogo de Aplicaciones (Catalog) como se muestra en la Figura 8.

HAM es el cerebro de la arquitectura y se encarga de buscar recursos disponibles, además de preparar, desplegar y controlar la infraestructura virtual. Mientras que el segundo elemento es el encargado de guardar y actualizar la información sobre la disponibilidad y el estado de los recursos de todo el sistema. HAM usa las herramientas del administrador de máquinas virtuales para realizar acciones como: consulta de recursos, despliegue de nuevas máquinas virtuales, conectar una máquina virtual con un acelerador de hardware, etc.

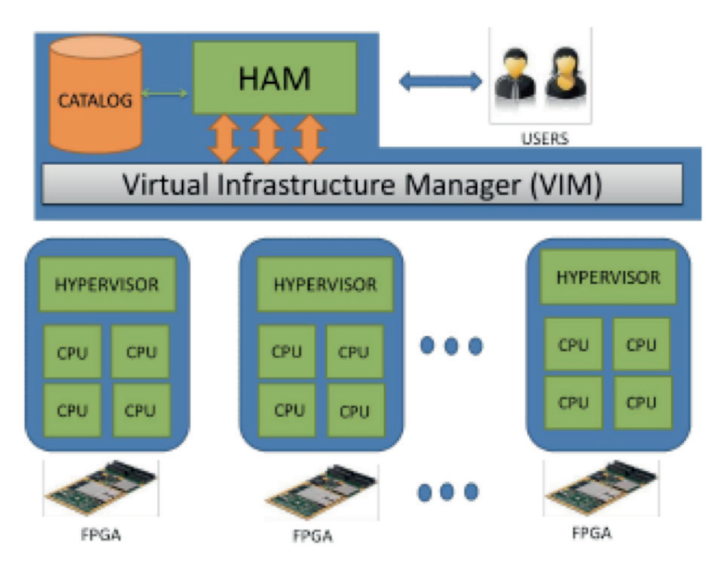


Figura 8. Modelo de arquitectura propuesta [37]

El Catálogo, por otra parte, almacena la información sobre los recursos, como por ejemplo número de procesadores de cada nodo, identificación de dispositivos de aceleración, ficheros de aplicaciones (Bitstreams) disponibles para ser usados por estos dispositivos. Además, mantiene actualizada esta información para su uso futuro.

Como se puede observar en la Figura 9, HAM está compuesto por cuatro controladores: *Controlador de Catálogo (CC), Controlador de Infraestructura Virtual (VIC), Controlador Bitstreams (BC) y finalmente un Controlador de Tareas (JMC)*, cada uno de ellos tiene una funcionalidad especifica que vamos a explicar a continuación.

El *Controlador de Infraestructura Virtual (VIC)* utiliza las funciones del administrador de máquinas virtuales para desplegar máquinas virtuales en los distintos nodos. Además, es el encargado de vincular cada máquina virtual con el hardware de aceleración respectivo.

El segundo elemento es el Controlador de Bitstreams (BC), este módulo ofrece dos funciones:

a) se encarga de transferir el fichero bitstream desde el servidor hasta el cliente a través de una conexión SSH.
b) es la encargada de configurar o programar el hardware FPGA usando las herramientas del fabricante.

Otro elemento fundamental es el *Controlador de Tareas (JMC)*, este módulo tiene la función de monitorizar el estado de una ejecución de una tarea, es decir, este módulo es una especie de demonio que se encarga de comunicar al sistema cuando la ejecución de una tarea ha cambiado de estado.

Finalmente, el último módulo es el encargado de interactuar con el *Catálogo (CC)*, este elemento se comunica con el Catálogo a través de operaciones de lectura y escritura. El Controlador del Catálogo mantiene actualizada la información de los recursos del sistema de tal manera que el administrador puede hacer uso de esta información cada vez que reciba una solicitud por parte del usuario o del propio sistema.

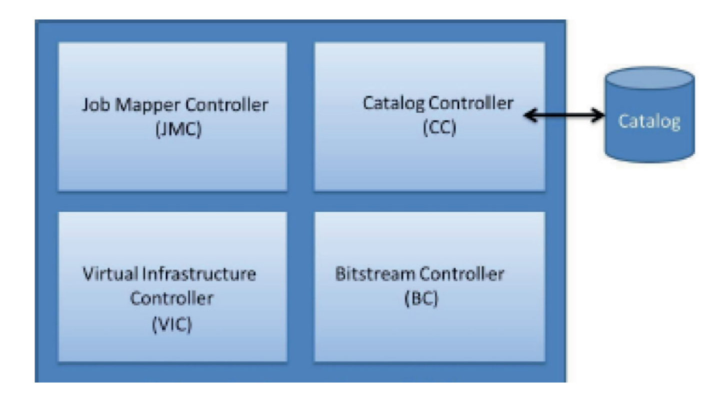


Figura 9. Gestor de aceleración de hardware [37]

A continuación se va a explicar los detalles de funcionamiento del sistema. Para ello, se ha dividido el proceso en tres fases. En la primera fase o puesta en marcha del sistema, el *Controlador de Catálogo* se pone en funcionamiento, consultando toda la información sobre los recursos disponibles (paso 1) en ese momento, como se puede ver en la Figura 10. Esta información es recolectada y enviada al Catálogo. Una vez con esta información (paso 2), el sistema se pone en modo de espera de peticiones. En este modo, el servidor se encuentra a la espera de llegada de una petición por parte del usuario o del propio sistema.

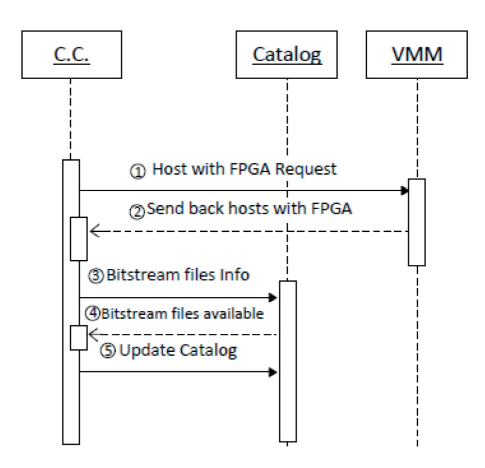


Figura 10. Controlador de bitstream [36]

La segunda fase inicia con una petición por parte del usuario. Esta petición se procesa por el Controlador de Infraestructura Virtual (VIC). Este controlador en primera instancia revisa si la solicitud esta correcta y completa (pasos 1 y 2). Los elementos que conforman esta solicitud son datos sobre la máquina virtual, como identificación, una imagen del sistema operativo que se pretende desplegar y la posibilidad de vincular esta máquina virtual con un hardware de aceleración (pasos 6, 7 y 8), como se puede observar en la Figura 11. Si toda esta información está completa y es correcta, se acepta la solicitud y se despliega el software virtual respectivo; en caso contrario, se muestra un mensaje de error.

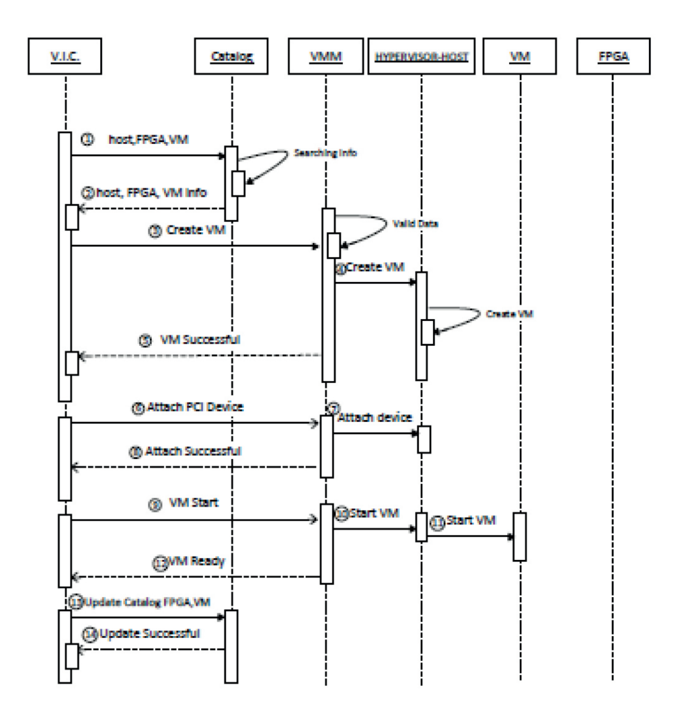


Figura 11. Controlador de infraestructura Virtual [36]

La arquitectura también pretende proporcionar la opción de que el administrador pueda agregar más imágenes o elementos adi-

cionales al Catálogo como recursos al sistema. Finalmente, una vez que tenemos a nuestra disposición el software virtual desplegado (máquinas virtuales) (pasos 9, 10 y 11) el siguiente paso es preparar el hardware de aceleración. Para ello el sistema realiza los pasos detallados a continuación: El Controlador de Catálogo revisa si dispone de la aplicación (pasos 1 y 2), los ficheros bitstream en el caso de FPGA o de la aplicación en el caso de GPU como se muestra en la Figura 12. Si se dispone de la aplicación, el Catálogo pone a disposición del Controlador de Bitstreams el fichero (bitstream). Este elemento es transferido y/o configurado para ser utilizado como acelerador (pasos 3, 4 y 5). Una vez que todo este proceso se haya realizado correctamente, el sistema está listo para ser explotado por el usuario (paso 6).

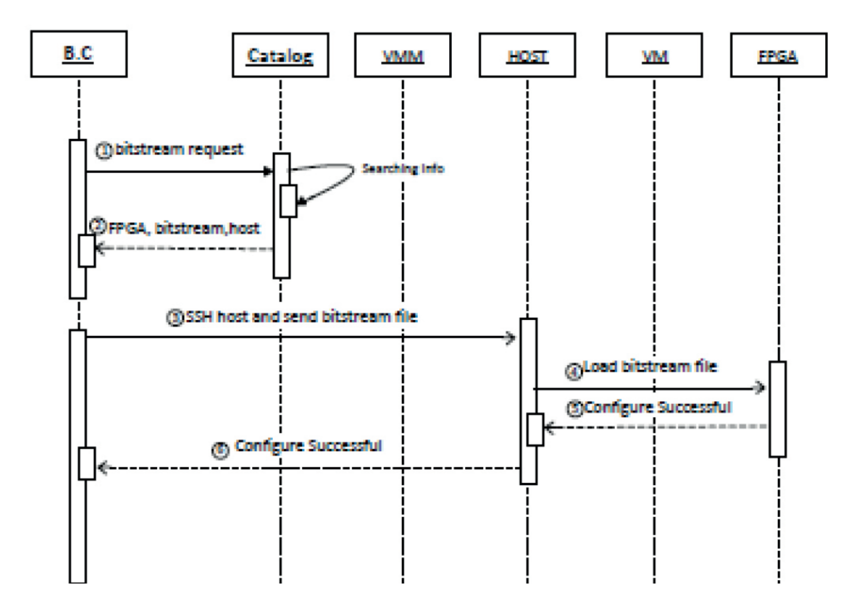


Figura 12. Configuración de bitstreams [36]

En el caso de que no se disponga de la aplicación en el Catálogo, el usuario tiene la posibilidad de subir su propio fichero al Catálogo. Por último, hay que mencionar que cuando el usuario haya terminado de usar los recursos virtuales, estos serán liberados del sistema.

B. Estrategia de gestión de recursos

La estrategia de asignación se implementa dentro del Controlador de Infraestructura Virtual, y se describe en la Figura 11. Se usa un método de aprendizaje supervisado para estimar el uso del tiempo de una solicitud de servicio particular, en base a los datos anteriores, es decir, los datos de entrenamiento.

El servicio VIC lleva a cabo dos etapas:

1. Implementa un modelo de predicción que aprovechan la información almacenado en el Catálogo. Tenga en cuenta que en nuestro sistema los usuarios no necesitan tener conocimiento previo de la duración real del servicio, por lo que el tiempo de finalización debe estimarse en este punto.
   Los datos de entrada se refieren a la complejidad del software (por ejemplo, la cantidad de los datos a procesar), y el resultado es el tiempo de procesamiento esperado en un recurso de procesamiento específico. Por lo tanto, este modelo se usa directamente por el sistema para hacer predicciones sobre el tiempo de ejecución esperado para las nuevas solicitudes enviadas.
2. El clasificador toma una decisión sobre qué tipo de se necesitan recursos para cumplir con la solicitud de servicio basado en la estimación de tiempo de ejecución total, incluyendo el momento de implementar la infraestructura virtual () y los parámetros de QoS establecidos en la solicitud de servicio. Por lo tanto, la solicitud de servicio puede se clasifica para ser ejecutado dentro de un elemento de procesamiento estándar (CPU) o con la cooperación de hardware de aceleración FPGA.

El objetivo principal de esta estrategia de asignación es evitar la violación de SLA utilizando elementos de proceso estándar y elementos de proceso muy rápidos como aceleradores trabajando juntos, al programar su uso en un "inteligente" camino.

Considere un ejemplo simple en donde una solicitud de servicio, como una Transformada de Fourier para 1024 Gb de datos, se envía al sistema. Supongamos también que no hay hardware de aceleración disponible. Entonces, el clasificador estima el tiempo necesario para completar esta solicitud usando un modelo de regresión lineal. Como posible escenario, el cliente define un tiempo para

usar el software como un servicio, pero el sistema puede estimar que el tiempo definido por el usuario no es suficiente para completar la solicitud si únicamente se ejecuta en CPUs convencionales. En ese caso, si la solicitud es aceptada, una violación de los niveles de servicio especificados se produciría. Sin embargo, cuando el sistema sabe que tiene hardware de aceleración, el servicio VIC clasificaría esta solicitud para programarse en el conjunto de recursos con hardware de aceleración CPU+FPGA. De esta manera, los niveles de servicio solicitados se cumplirían. Como efecto secundario, y en el caso particular de FPGA como aceleradores, existen ventajas adicionales, como el menor consumo de energía.

Una vez que la solicitud de servicio ha sido clasificada para su ejecución, el JMC debe buscar un espacio libre dentro del conjunto adecuado de recursos para ejecutar la solicitud de servicio, y dentro del tiempo intervalo establecido en la solicitud de servicio. Para tal fin, se pueden aplicar varios criterios de búsqueda, como First-Fit o incluso del que mejor ajuste (es decir, programar la solicitud en el primer espacio libre el tiempo suficiente para cumplir la solicitud). La solicitud dentro del intervalo de tiempo especificado por el usuario, o en la ranura cuya duración mejor se ajusta al tiempo estimado de ejecución del servicio. Criterios adicionales pueden incluye tratar de minimizar el número de ejecuciones VMs para minimizar el consumo de energía.

En esta propuesta se considera el algoritmo First-Fit como criterio de asignación de recursos. Esto implica el uso de la estimación calculada en el módulo JMC. Esto se hace para elegir el recurso más adecuado para asistir a la solicitud actual, evitando así el desperdicio de uso de recursos (es decir, cuando el usuario ha especificado un tiempo marco para la ejecución del servicio mucho más grande que el real tiempo de ejecución esperado). Además, el tiempo de implementación también son considerados en esta propuesta. Este tiempo considera las acciones de iniciar la máquina virtual con los parámetros tales como imágenes, memoria, software, y controladores previamente definidos (ver Figura 14).

5. Evaluación

El enfoque descrito en secciones anteriores ha sido implementado en un entorno privado real heterogéneo en la Universidad de

Castilla La Mancha (España) [37], en el cual se ha incluido una FPGA. El objetivo de esta evaluación es realizar una prueba de concepto de la arquitectura propuesta. Además, los experimentos también se han diseñado para garantizar la repetitividad.

Para comparar el impacto de agregar hardware como dispositivos de aceleración de aplicaciones hemos dividido estos experimentos en dos escenarios. Por un lado, un escenario estándar de Cloud privada con solo CPU como procesador, este constituirá nuestra infraestructura de control y se usa como referencia para la comparación.

El otro escenario consiste en incluir una FPGA como un elemento de procesamiento de aceleración en uno de los nodos (ver Figura 13).

En ambos escenarios, los clientes enviarán solicitudes para un servicio particular. Para este casi se ha seleccionado la Transformada Discreta de Fourier.

Para evaluar el desempeño se ha considerado la tasa de cumplimiento del acuerdo de nivel de servicio (SLA). Así, se contabiliza el número de cumplimientos de SLA. Por el contrario, una violación de SLA ocurre cuando el sistema no puede completar el servicio de software dentro de los requisitos de tiempo establecidos en la solicitud del usuario. Por lo tanto, SLA cumplidos son aquellos que se completan sin sufrir una violación.

La siguiente subsección entra en más detalles sobre banco de pruebas y la carga de trabajo.

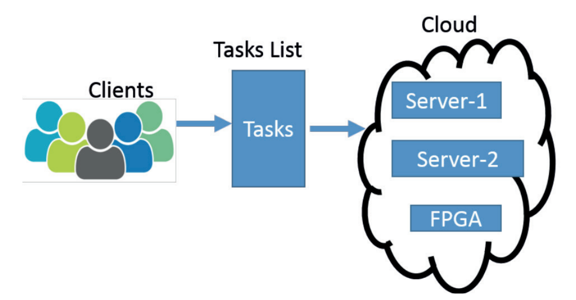


Figura 13. Entorno de pruebas

A. Carga de trabajo

La infraestructura del banco de pruebas se divide en dos actores: el proveedor de servicios de Cloud y los clientes. Respecto al

proveedor, el módulo HAM ha sido desarrollado sobre el administrador de infraestructura virtual (VMM), OpenNebula [7]. Además, se ha configurado KVM [23] como hypervisor.

El hardware utilizado consiste en dos diferentes servidores. Uno de ellos es un Intel Core i5 con 6 Gb de la memoria RAM, con un Xilinx FPGA Virtex 6 entrelazado a su bus PCIe. La FPGA Virtex 6 es un ML605 Evaluation Kit, basado en Xilinx XC6VLX-240T-1FFG1156 [24].

El otro servidor consiste en un Intel Quad Core con 6Gb de RAM memoria sin FPGA. El servicio ofrecido es una emulación de lo La Transformada Discreta de Fourier. La transformada Discreta de Fourier (DFT) es una herramienta importante en muchas ramas de la ciencia y la ingeniería [25], como en el procesamiento de señales digitales o procesamiento de imágenes. El DFT es también usado para resolver eficientemente ecuaciones diferenciales parciales, y para realizar otras operaciones como convoluciones [26]. Para la evaluación de la propuesta se usa el algoritmo de Cooley-Tukey [27]. Se ha elegido este algoritmo para reducir la complejidad computacional de O (N2) a O (Nlog2N) [28], y la regularidad del algoritmo lo hace adecuado para la implementación en hardware. Los detalles sobre el algoritmo completo se detallan en [28].

Para emular la parte de clientes solicitando el servicio con diferentes necesidades informáticas, el algoritmo se ejecuta un cierto número de veces sobre el la misma cantidad de datos. Entonces, el parámetro asociado con cada servicio solicitado indica la complejidad del trabajo.

En la emulación, los clientes envían solicitudes de servicio DFT al proveedor. En esta solicitud se definen dos parámetros. El primero se refiere al tiempo de uso, y contiene el tiempo de inicio y finalización de la tarea. Por ejemplo, tstart = 12: 00: 00 y tend = 12: 35: 00. Esto significa que el servicio debe ser provisto dentro de este periodo de tiempo. Sin embargo, estos parámetros de entrada son no obligatorios. Si no se proporciona tstart, el servicio será considerado como una solicitud inmediata y el tiempo del sistema se asignará a tstart (es decir, tstart = tiempo). Si no se proporciona tend, el servicio será considerado como uno de mejor esfuerzo, sin ninguna restricción de tiempo de finalización. En este caso, cualquier recurso libre se puede asignar para la ejecución del servicio.

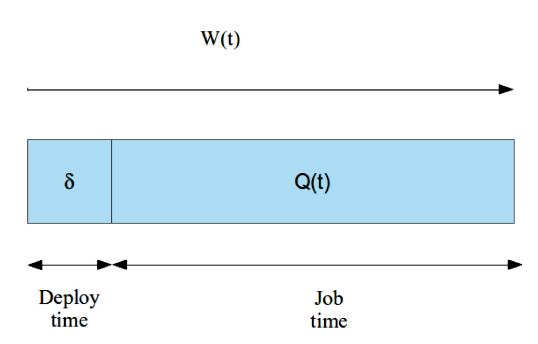


Figura 14. Definición de una tarea de usuario

En el segundo parámetro se define el nombre del servicio como tal y la complejidad computacional del servicio definida como el número de DFTs para calcular sobre un cierto conjunto de datos.

Se utiliza una distribución de Poisson como tasa de llegada de solicitudes, con una tasa $\lambda = 1/1$, es decir, se genera una solicitud de servicio cada minuto. Los parámetros de cada solicitud de servicio se generan de forma aleatoria. En particular, cada cliente generará una solicitud de servicio con N aleatoriamente entre 30000 y 300000. La hora de inicio (tstart) se ha establecido en la hora actual cuando se genera cada solicitud. Se debe tener en cuenta que la hora de inicio se puede configurar en cualquier momento después del tiempo de envío.

Para la hora de finalización, se han considerado dos tipos de solicitudes de servicio. Por un lado, los "clientes normales" que envían solicitudes cuyo tiempo de finalización (tend) se calculará como una distribución normal con $\mu = 9$ minutos y $\sigma = 0.5 * \mu$. Por otro lado, los "clientes en tiempo real" exhiben tiempos de uso mucho más exigentes. Su tiempo de finalización se establecerá según una distribución normal con $\mu = 0.9$ minutos y $\sigma = 0.5 * \mu$. Recuerde que el sistema rechazará aquellas solicitudes de servicio que no se pueden atender dentro del marco de tiempo especificado con los recursos disponibles. Por lo tanto, se espera que los "clientes en tiempo real" necesiten hardware de aceleración para ser atendido adecuadamente.

En la siguiente sección, se analiza el impacto de agregar elementos de FPGA como hardware de aceleración y su administración con las técnicas propuestas.

B. Resultados

La arquitectura y los algoritmos descritos en las secciones anteriores se han probado generando una carga de trabajo compuesta por 60 clientes. La misma carga de trabajo tiene sido aplicado a los dos escenarios, servidores solo con CPUs y servidores con FPGA. Además, cada experimento ha sido repetido tres veces, y los resultados que se muestran aquí son los valores promediados.

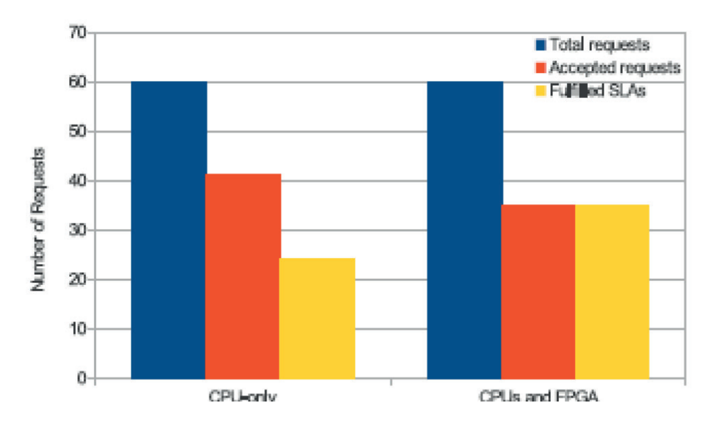


Figura 15. Total de peticiones que cumplen SLA

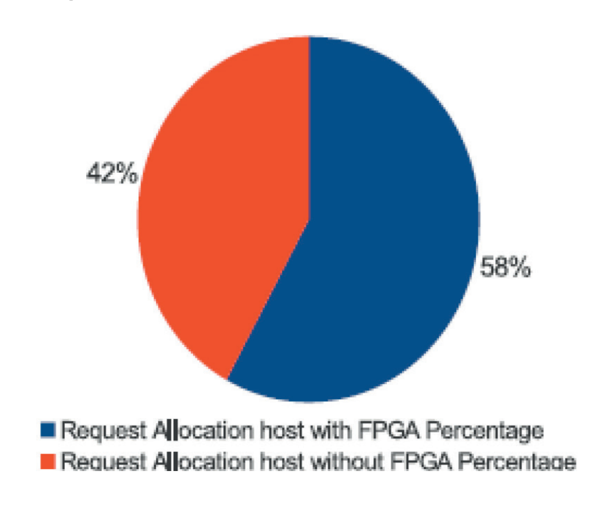


Figura 16. Distribución de peticiones

En cada experimento, tanto "clientes normales" como Los "clientes en tiempo real" se envían solicitudes de servicio al sistema. Como los nodos tienen ocho núcleos y una FPGA la razón de solicitudes de tiempo real es de 1/7; es decir que por cada siete solicitudes normales se genera 1 de tiempo real. Las solicitudes han sido mezcladas al azar siguiendo esta proporción.

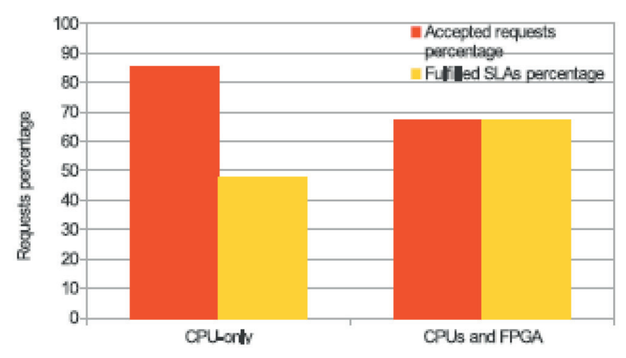


Figura 17. Total de peticiones que cumplen SLA

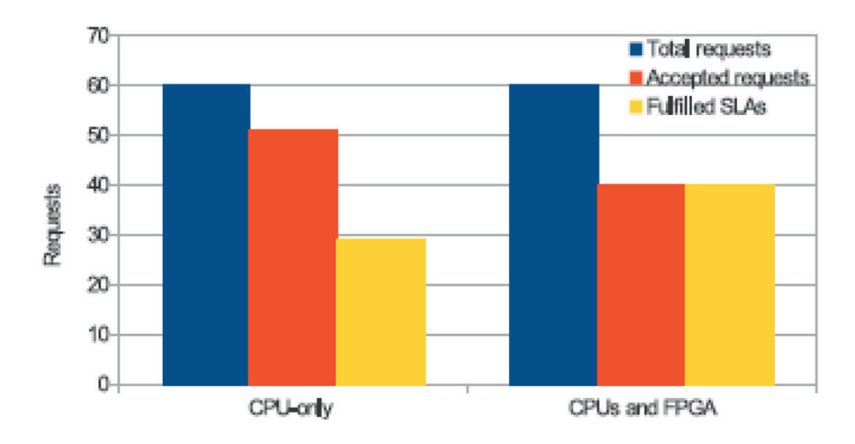


Figura 18. Total de peticiones que cumplen con SLA

Como se muestra en la Figura 18, el número de solicitudes aceptadas para el escenario en el que se utilizan solo CPUs es 51, y el número de solicitudes sin violación de SLA es 29. Por lo tanto, resulta

que el 57% de las solicitudes han sido atendidas con éxito. En el caso del escenario de CPU y FPGA, el número de solicitudes aceptadas es de 40, mientras que el 100% de las solicitudes cumplen el SLA. A pesar del hecho de que el número de solicitudes aceptadas es ligeramente menor que en el escenario en el cual únicamente se usa CPU, el porcentaje de solicitudes que se sirven con éxito es del 100%.

Finalmente, la proporción de solicitudes atendidas por cada clase de recursos se muestra en la Figura 16. Se puede ver que se ha asignado el 67% de las solicitudes aceptadas al FPGA. En resumen, la fiabilidad de la arquitectura y las técnicas propuestas al proporcionar garantías de QoS se refleja en los resultados. Estos son resultados prometedores, especialmente porque solo hay un FPGA disponible como acelerador.

Para el segundo escenario los resultados son similares como se puede observar en la Figura 17, en el caso que se usa la combinación de CPU y FPGA todas las solicitudes aceptadas por el sistema son cumplidas. Finalmente, en la Figura 19 se puede observar que la introducción de solicitud de mayor demanda hace que el sistema distribuya de mejor manera las peticiones a los recursos para este caso se concentran más peticiones para los recursos que tienen CPU y FPGA (67%). Así, el sistema puede lograr distribuir de manera más eficiente la carga, aprovechando mejor las características de los recursos.

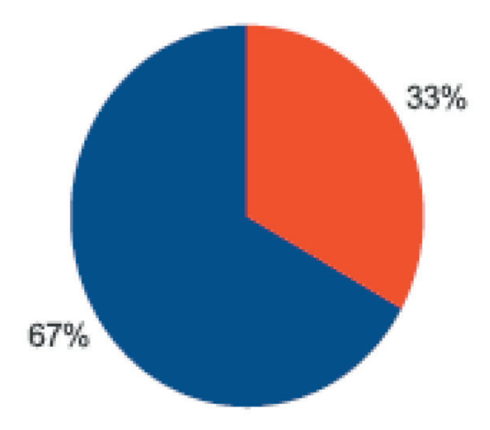


Figura 19. Distribución de peticiones

6. Conclusiones

Una infraestructura heterogénea Cloud en el que los aceleradores tales como FPGA y GPU se utilicen para reducir el consumo de energía puede ser una excelente forma de utilizar los recursos de una manera más eficiente. Por lo tanto, en este trabajo nos hemos enfocado en una arquitectura que soporte agregar un elemento de computación de hardware como acelerador y que pueda ser ofrecida al usuario como servicio. El trabajo futuro consiste en la implementación de la arquitectura propuesta. Además en un futuro se propone enfocar todos los esfuerzos en el diseño de una herramienta para planificación del uso efectivo de los recursos en el tiempo. Adicionalmente, exploraremos la factibilidad de usar la tecnología que ofrece Intel (SR-IOV) [15], para que múltiples máquinas virtuales puedan compartir el hardware de aceleración al mismo tiempo, junto con las funcionalidades que ofrece la reconfiguración dinámica y reconfiguración parcial de FPGAs [5].

Agradecimientos

Este trabajo se ha desarrollado con la colaboración del Instituto de Investigación en Informática de Albacete (España) y la Universidad Politécnica Salesiana (Quito-Ecuador). Además, es muy importante agradecer a la Secretaría Nacional de Ciencia y Tecnología (Senescyt) por el apoyo a proyectos de investigación. Finalmente, un agradecimiento especial al grupo ARCO de la Universidad de Castilla La Mancha, por el uso de su infraestructura y el soporte dado durante todo el desarrollo de este proyecto de investigación.

Referencias

[1] L. Wang *et al.*, “Cloud computing: a perspective study,” *New Gener. Comput.*, vol. 28, no. 2, pp. 137–146, 2010.

[2] R. Buyya, C. S. Yeo, and S. Venugopal, “Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities,” in *High Performance Computing and Communications, 2008. HPCC’08. 10th IEEE International Conference on*, 2008, pp. 5–13.

[3] T. Erl, R. Puttini, and Z. Mahmood, *Cloud computing: concepts, technology & architecture*. Pearson Education, 2013.

[4] L. Wu, S. K. Garg, and R. Buyya, "SLA-based admission control for a Software-as-a-Service provider in Cloud computing environments," *J. Comput. Syst. Sci.*, vol. 78, no. 5, pp. 1280–1299, 2012.
[5] J. Shuja *et al.*, "Survey of techniques and architectures for designing energy-efficient data centers," *IEEE Syst. J.*, vol. 10, no. 2, pp. 507–519, 2016.
[6] J. Glanz, "Power, pollution and the internet," *New York Times*, vol. 22, 2012.
[7] A. Shehabi *et al.*, "United states data center energy usage report," 2016.
[8] K. Freund, "Amazon and xilinx deliver new fpga solutions." Forbes, 2017.
[9] R. Uhlig *et al.*, "Intel virtualization technology," *Computer (Long. Beach. Calif).*, vol. 38, no. 5, pp. 48–56, 2005.
[10] M. Fenn, M. Murphy, J. Martin, and S. Goasguen, "An evaluation of KVM for use in cloud computing," in *Proc. 2nd International Conference on the Virtual Computing Initiative, RTP, NC, USA*, 2008.
[11] D. Abramson *et al.*, "Intel Virtualization Technology for Directed I/O.," *Intel Technol. J.*, vol. 10, no. 3, 2006.
[12] I. AMD and O. Virtualization, "Technology (IOMMU) Specification," 2007.
[13] O. Sefraoui, M. Aissaoui, and M. Eleuldj, "OpenStack: toward an open-source solution for cloud computing," *Int. J. Comput. Appl.*, vol. 55, no. 3, 2012.
[14] B. Sotomayor, R. S. Montero, I. M. Llorente, I. Foster, and others, "Capacity leasing in cloud systems using the opennebula engine," in *Workshop on Cloud Computing and its Applications*, 2008, vol. 3, p. 22.
[15] P. Mell, T. Grance, and others, "The NIST definition of cloud computing," 2011.
[16] M. Sachdeva, P. Rana, R. Kapoor, and M. Shahid, "Cloud computing-pay as you go technology," in *Proceedings of the 5th National Conference on Computing for National Development*, 2011, vol. 161.
[17] R. Buyya, S. K. Garg, and R. N. Calheiros, "SLA-oriented resource provisioning for cloud computing: Challenges, architecture, and solutions," in *Cloud and Service Computing (CSC), 2011 International Conference on*, 2011, pp. 1–10.
[18] J.-H. Morin, J. Aubert, and B. Gateau, "Towards cloud computing SLA risk management: issues and challenges," in *System Science (HICSS), 2012 45th Hawaii International Conference on*, 2012, pp. 5509–5514.
[19] M. Alhamad, T. Dillon, and E. Chang, "Conceptual SLA framework for cloud computing," in *Digital Ecosystems and Technologies (DEST), 2010 4th IEEE International Conference on*, 2010, pp. 606–610.
[20] L. Tomás and J. Tordsson, "Improving cloud infrastructure utilization through overbooking," in *Proceedings of the 2013 ACM Cloud and Autonomic Computing Conference*, 2013, p. 5.
[21] M. Balakrishnan *et al.*, "Tango: Distributed data structures over a shared log," in *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*, 2013, pp. 325–340.
[22] A. Berl *et al.*, "Energy-efficient cloud computing," *Comput. J.*, vol. 53, no. 7, pp. 1045–1051, 2010.

[23] T. V. T. Duy, Y. Sato, and Y. Inoguchi, “Performance evaluation of a green scheduling algorithm for energy savings in cloud computing,” in *Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), 2010 IEEE International Symposium on*, 2010, pp. 1–8.

[24] M. Mezmaz *et al.*, “A parallel bi-objective hybrid metaheuristic for energy-aware scheduling for cloud computing systems,” *J. Parallel Distrib. Comput.*, vol. 71, no. 11, pp. 1497–1508, 2011.

[25] K. Boloor, R. Chirkova, Y. Viniotis, and T. Salo, “Dynamic request allocation and scheduling for context aware applications subject to a percentile response time SLA in a distributed cloud,” in *Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on*, 2010, pp. 464–472.

[26] G. A. Constantinides, “FPGAs in the cloud,” in *Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, 2017, p. 167.

[27] E. Billauer, “Xillybus.” Haifa, 2017.

[28] S. Mohl, “The Mitrion-C programming language,” *Mitrionics Inc*, 2006.

[29] M. Jacobsen and R. Kastner, “RIFFA 2.0: A reusable integration framework for FPGA accelerators,” in *2013 23rd International Conference on Field programmable Logic and Applications*, 2013, pp. 1–8.

[30] M. Oikawa, A. Kawai, K. Nomura, K. Yasuoka, K. Yoshikawa, and T. Narumi, “DS-CUDA: A middleware to use many GPUs in the cloud environment,” in *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion:*, 2012, pp. 1207–1214.

[31] J. E. Stone, D. Gohara, and G. Shi, “OpenCL: A parallel programming standard for heterogeneous computing systems,” *Comput. Sci. Eng.*, vol. 12, no. 3, pp. 66–73, 2010.

[32] K. Paranjape, S. Hebert, and B. Masson, “Heterogeneous computing in the cloud: Crunching big data and democratizing HPC access for the life sciences,” *Intel Corp.*, 2010.

[33] P. Grigoras, M. Tottenham, X. Niu, J. G. F. Coutinho, and W. Luk, “Elastic Management of Reconfigurable Accelerators,” in *2014 IEEE International Symposium on Parallel and Distributed Processing with Applications*, 2014, pp. 174–181.

[34] I. Arpaci, “Understanding and predicting students’ intention to use mobile cloud storage services,” *Comput. Human Behav.*, vol. 58, pp. 150–157, 2016.

[35] R. Moreno-Vozmediano, Montero, and I. M. Llorente, “IaaS Cloud Architecture: From Virtualized Datacenters to Federated Cloud Infrastructures,” *Computer (Long. Beach. Calif).*, vol. 45, no. 12, pp. 65–72, Dec. 2012.

[36] J. Proaño, C. Carrión, and M. B. Caminero, “An Open-Source Framework for Integrating Heterogeneous Resources in Private Clouds.,” in *CLOSER*, 2014, pp. 129–134.

[37] J. P. Orellana, M. B. Caminero, and C. Carrión, “On the provision of SaaS-level quality of service within heterogeneous private clouds,” in *Utility and*

Cloud Computing (UCC), 2014 IEEE/ACM 7th International Conference on, 2014, pp. 146–155.

[38] J. Proaño, C. Carrión, and M. B. Caminero, "Towards a green, QoS-enabled heterogeneous cloud infrastructure," in *Parallel and Distributed Processing Symposium Workshops, 2016 IEEE International*, 2016, pp. 7–16.

[39] A. Putnam *et al.*, "A reconfigurable fabric for accelerating large-scale datacenter services," in *2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)*, 2014, pp. 13–24.

[40] Y.-M. Choi and H. K.-H. So, "Map-reduce processing of k-means algorithm with FPGA-accelerated computer cluster," in *Application-specific Systems, Architectures and Processors (ASAP), 2014 IEEE 25th International Conference on*, 2014, pp. 9–16.

[41] K. Neshatpour, M. Malik, and H. Homayoun, "Accelerating machine learning kernel in hadoop using fpgas," in *Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on*, 2015, pp. 1151–1154.

Sobre los autores



Julio Proaño es profesor titular la Universidad Politécnica Salesiana (Quito-Ecuador). Es PhD en Tecnologías Informáticas Avanzadas en la Universidad de Castilla (La Mancha, España. Recibió su título de grado en Ingeniería Electrónica en la Universidad Salesiana Politécnica (Cuenca-Ecuador) en 2008 y su Maestría en Informática en la Universidad de Castilla (La Mancha, España en 2012. Sus intereses actuales de investigación son los Gestión de recursos heterogéneos (FPGAS) y QoS con SLA en entornos de Cloud Computing. Él es un miembro del IEEE.



Maria Blanca Caminero es Profesora Asociada en Arquitectura de Computadores y Tecnología en Computación Departamento de Sistemas La Universidad de Castilla La Mancha. Ella tiene un PhD. en Informática de la Universidad de Castilla La Mancha, y sus intereses actuales de investigación son Soporte de QoS y meta-programación en Grids y Nubes. Ella es miembro del IEEE



Carmen Carrión es Profesora Asociada en Arquitectura de Computadora y Tecnología en el Departamento de Sistemas Computacionales de la Universidad de Castilla La Mancha. Tiene un doctorado en Física de la Universidad de Cantabria, y sus intereses incluyen meta-programación y QoS en Grids and Cloud.

CAPÍTULO VI

Despliegue óptimo de redes PON, usando mapas reales, para un gran número de usuarios con diversas demandas de ancho de banda

Optimal deployment of PON, using real maps, for a large number of users with heterogeneous bandwidth demands

Germán Arévalo*, Roberto Gaudino§
Universidad Politécnica Salesiana Ecuador, Politécnico Di Torino Italia
*Email: garevalo@ups.edu.ec
§Email: roberto.gaudino@polito.it

Resumen

El problema relacionado con la búsqueda de topologías óptimas para despliegue de PON en una ciudad, para servir a un determinado número de usuarios, es en esencia un problema del tipo Optimización del Árbol de Steiner, el cual es un bien conocido problema del tipo NP-Hard. Así, para proponer soluciones para este tipo de problemas se deben emplear procesos de tipo heurístico. Por otro lado, se tiene el problema de seleccionar la tecnología PON más apropiada para satisfacer las demandas de ancho de banda de los usuarios, con escalabilidad e inversión redituable. En el presente artículo se detalla los aspectos técnicos y resultados obtenidos a través de un algoritmo para el despliegue óptimo de redes PON en grandes regiones urbanas caracterizadas por un gran número de usuarios potenciales de tipo residencial y corporativo que demandan diversos anchos de banda. El estudio presentado incluye una comparación del costo de implementación de distintos tipos y estándares de redes PON.

Palabras clave

PON, estudio tecno-económico, optimización, algoritmo, heurísticas.

Abstract

The problem related with the optimal search of topologies for the deployment of PON in an urban region, for servicing a given number of users, is essentially a Steiner-tree problem, which is a well-known NP-Hard problem. Therefore, in order to find solutions for this type of problems it is necessary the use of heuristic methods. On the other hand, there is the problem related with the selection of the most suitable PON technology, in order to satisfy the users' bandwidth demands in a scalable and profitable way. This paper details the technical aspects and the results obtained through the use of an algorithm which finds the optimal PON deployment in large urban regions with very large number of users with heterogeneous bitrate demands. The study covers the costs comparison among different types of PON standards.

Keywords

PON, techno-economics, optimization, algorithm, heuristics.

1. Introducción

Debido a los múltiples problemas confrontados en el desarrollo de redes ópticas de acceso de nueva generación (NG), el desarrollo de soluciones que viabilicen el despliegue comercialmente viable de este tipo de redes, a constituido un tópico de investigación muy popular en los últimos años, especialmente el estudio de las redes ópticas pasivas (PON) de nueva generación [1].

Particularmente, el estudio tecno-económico de las redes PON de nueva generación, mismas que incluyen en la tecnología de la red técnicas de transmisión y compartición del canal más complejas y avanzadas que las empleadas por los estándares actualmente existentes, como la multiplexación por división de longitud de onda (WDM) incluyendo la división de longitud de onda ultra-densa (UDWDM), ha recibido especial atención por parte de varios investigadores [2][3][4][5]. Las estrategias para la optimización este tipo de redes se basan en el estudio del despliegue de la red de distribución óptica (ODN) con el menor costo posible [6] [7][8], o en el diseño del hardware de la red [9][10].

En el año 2015 ya se publicó el estándar UIT para las redes ópticas pasivas de nueva generación, el Next-Generation Passive Optical Network - Version 2 (NGPON2) [11].

Se han propuesto varias alternativas para la topología de una red óptica de acceso, sin embargo la topología tipo árbol de PON es la única actualmente existente y comercialmente implementada en el mundo y que ha sido estandarizada para este tipo de redes [12]. La topología tipo árbol se basa exclusivamente en el empleo de divisores de potencia ópticos (splitters) en cascada con varias etapas, aunque en la práctica las PON se implementan hasta con dos niveles de splitting. En la Figura 1 se ilustra la topología lógica de una PON, donde puede observarse que la red parte de un transceptor óptico denominado “terminal de línea óptica” (OLT), ubicado en la “oficina central” (CO) del proveedor de servicios. A partir de ahí, a través de una fibra óptica (OF) denominada “fibra de alimentación” (Feeder) se arriba a un splitter remoto generalmente ubicado a varios kilómetros de distancia, dentro de un gabinete que generalmente se encuentra en una esquina sobre la vereda o en algún lugar de acceso público, al que se le denomina “gabinete de calle primario” (PSC). Desde ahí, a través de cables ópticos de distribución de menor longitud se llega o directamente al usuario final o a un segundo splitter para luego con cables ópticos de corta distancia (si el segundo splitter está en la entrada de un edificio, estos últimos cables generalmente con cables para interiores). En el extremo del usuario final se encuentra un equipo de transmisión-recepción denominado la “unidad óptica de red” (ONU). La red óptica que parte desde la OLT y llega a la ONU constituye la ODN.

Sin embargo, en la práctica, para cubrir a un gran número de usuarios se realiza un despliegue simultáneo de múltiples redes PON, desde varias CO con muchas OLT (colocadas en racks de OLT), compartiendo de manera parcial las rutas y cables que se despliegan y en consecuencia la infraestructura de la instalación (i.e. ductos, zanjas, cables subterráneos y aéreos, gabinetes de calle, etc.). En la Figura 2 se puede observar la topología física típica de una implementación real de múltiples PON en una región urbana.

Oficina Central (CO)
Racks OLT
Feeder
PSC
Cables ópticos
de distribución
Edificio residencial
PON
SSC
ONU
ONU
Anillo óptico
metropolitano
Edificio corporativo
PON
SSC
ONU
Cable de
interiores
Oficina Central (CO)
Racks OLT
Cable óptico de
alimentación
(Feeder)
Edificio residencial
PON
Cables ópticos
de interiores
ONU
ONU
PSC
SSC
PON
ONU
ONU
Cables ópticos
de distribución

Figura 1. Topología lógica y nomenclatura de una PON

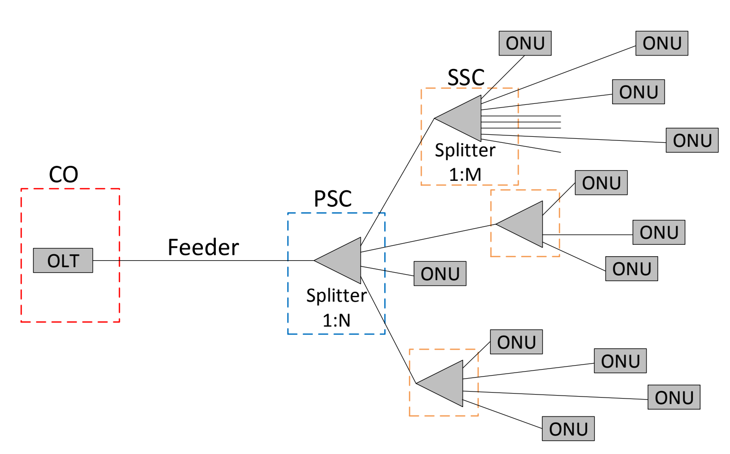


Figura 2. Topología física típica de un despliegue de múltiples PON en una región urbana con muchos usuarios

Dada la complejidad que implica la instalación de este tipo de topologías, el estudio de la planificación para el depliegue óptimo de PON es de mucho interés para los operadores [13].

El resto del presente artículo está organizado como sigue: La sección 2 describe los parámetros generales, tanto de los escenarios urbanos a ser considerados, como los costos referenciales y parámetros técnicos de los distintos tipos de PON que se cubren en este estudio. En la sección 3 se presentan los detalles de la formulación del problema de optimización que modela la búsqueda del despliegue de menor costo para múltiples PON en un escenario urbano extenso. La sección 4 explica los detalles del algoritmo matemático desarrollado para la solución del problema. En la sección 5 se muestran los resultados más relevantes de la investigación y, finalmente la sección 6 concluye el artículo.

2. Parámetros y variables del problema

A. *Consideraciones generales*

El presente estudio se enfoca en el despliegue de redes PON para proveer fibra óptica al hogar (FTTH) a través de un óptimo dimensionamiento y planeación de la ODN.

Para tal fin se consideran tres tipos de PON estandarizadas: (i) la GPON, (ii) la XGPON, [19] mismas que pueden conectar hasta 64 usuarios empleando TDM/TDMA como técnicas de compartición del canal para bajada "downstream - DS" y subida "upstream - US", respectivamente, y (iii) el nuevo estándar NGPON2 [11], introducido por primera vez en el 2013 y cuya última revisión se publicó en el 2015, misma que es una red híbrida TDM/WDM con entre cuatro a ocho canales WDM densos (DWDM) tanto para DS como para US. NGPON2 no reemplaza los splitters por multiplexores, sino los mantiene con el fin de proveer compatibilidad con la ODN de estándares anteriores de redes PON.

Algunos trabajos de información relevantes proponen soluciones de hardware de bajo costo para NGPON2, como el reportado en [20]. También usamos en nuestro análisis una propuesta de red de

nueva generación a largo plazo denominada UDWDM-PON (PON con multiplexación ultra-densa de longitudes de onda) [21].

El escenario que se emplea para el desarrollo del presente estudio es el de una región urbana (i.e. una zona dentro de una ciudad o urbe) en la cual existe una gran densidad de habitantes y de entre ellos un gran número de usuarios (en el orden de cientos de miles) demandan servicios de acceso a Internet, con distintos anchos de banda, a través de conexiones FTTH. Tal número de usuarios no podrían ser cubiertos por una sola CO, de modo que se considera la presencia de varias CO en la región, cada una de ellas cubriendo una subregión de hasta aproximadamente 20000 usuarios a través de varios cientos de OLT en sus racks (i.e. cada OLT significa una PON individual a ser desplegada en la ciudad). Se asume que las CO están interconectadas entre sí, por ejemplo, a través de un anillo óptico metropolitano, como el que se observa en la Figura 2.

B. Costos referenciales

A través de interacción directa con proveedores de servicios y vendedores de equipos, hemos obtenido la información de costos referenciales que se detallan a continuación, en las Tablas 1, 2 y 3.

Tabla 1. Costos de cables ópticos y trabajos de instalación

Componente	Costo ($)
Feeder Cable, 2 fibras /km	600
Feeder Cable, 4 fibras /km	800
Feeder Cable, 6 fibras /km	1000
Feeder Cable, 12 fibras /km	1500
Feeder Cable, 24 fibras /km	2000
Feeder Cable, 48 fibras /km	2500
Feeder Cable, 64 fibras /km	3000
Feeder Cable, 96 fibras /km	3500
Feeder Cable, 140 fibras /km	3700
Feeder Cable, 280 fibras /km	4000

Componente	Costo ($)
Distribution Cable /km	2000
Cable de interiores /usuario	50
Zanjado y recapeo/km	30000
Ductos/km	10000
Empalmes/unidad	10
Pozos de revisión/unidad	500

Tabla 2. Costo de gabinetes

Componente	Costo ($)
Caja de empalmado 140 OF	500
Caja de empalmado 48 OF	400
Caja de empalmado 16 OF	350
Caja de empalmado 8 OF	300
1:64 splitter	120
1:32 splitter	70
1:16 splitter	45
1:8 splitter	28
1:4 splitter	24
1:2 splitter	20
Gabinete (incluida instalación)	1600

Tabla 3. Costos del hardware PON

Componente	Costo ($)
OLT chasis - GPON (103 usuarios)	16000
OLT chasis - XGPON (103 usuarios)	28000
OLT chasis - NGPON2 (103 usuarios)	50000
OLT chasis - UDWDM PON (103 usuarios)	85000

Componente	Costo ($)
OLT tarjeta - 8xGPON	9000
OLT tarjeta - 8xXGPON	15000
OLT tarjeta - 8xNGPON2	25000
OLT tarjeta - 8xUDWDW-PON	40000
ONU residencial - GPON	100
ONU residencial - XGPON	350
ONU residencial - NGPON2	600
ONU residencial - UDWDM PON	1100
ONU corporativa - GPON	350
ONU corporativa - XGPON	600
ONU corporativa - NGPON2	1100
ONU corporativa - UDWDM PON	2200
Fusión / por empalme	10
Instalación de la OLT	2000
ODF (para cada rack OLT)	3500

C. Parámetros de la red

En la Tabla 4 se especifican los parámetros generales de cada tecnología PON que se considera en este estudio. El enfoque de la investigación es en la demanda de ancho de banda de bajada, DS, por parte de los usuarios, dado que el tráfico DS por parte de los usuarios es mucho mayor que el US. En el caso de UDWDM PON, se emplean los parámetros de la propuesta de Rohde reportada en [22], con la única variación de que en lugar de los 1024 usuarios a 1 Gb/s que se pueden conectar con la propuesta de Rohde, en este trabajo se considera un número más conservativo: 256 usuarios por cada OLT.

Además, se consideran los siguientes:

- Tipo de OF: UIT G652 D (monomodo estándar)

- Atenuación en los splitters: α =3,5*$\log_2(k)$ dB [99] (donde k es el número de ramales de salida del splitter).
- Bitrate referencial (para normalización): BR_{ref} = 10 Gb/s.

En lo que respecta a las demandas de ancho de banda, o bitrate (BR), consideramos dos tipos de usuarios: residenciales y corporativos (donde el número de corporativos representan apenas el 2%, con es lo usual). Las demandas se consideran del tipo "ancho de banda mínimo garantizado" y para simular un escenario realista, estas son seleccionadas aleatoriamente dentro de los seis posibles rangos especificados en la Tabla 5.

Tabla 4. Parámetros de las tecnologías PON consideradas

Componente	Tecnología PON			
	GPON	XGPON	NGPON2	UDWDM PON
Max. longitud de enlace [km]	40	40	40	100
Max. atenuación de ODN[dB]	35	35	35	43
Usuarios por cada OLT	64	64	64	256
Num. de longitudes de onda	1	1	4	256
DS bitrate por OLT [Gb/s]	2.5	10	40	256

Tabla 5. Escenarios de demanda de bitrate

Escenario	Bitrate (Mb/s) por tipo de usuario	
	Residencial	Corporativo
1	10 - 50	100 – 500
2	50 – 100	500 - 1000
3	100 – 400	1000 - 2500
4	100 - 1000	1000 - 10000
5	500 - 2500	2500 - 10000
6	1000 - 2500	5000 - 40000

3. Formulación del problema

A. *Definición de conjuntos y variables*

Una región urbana donde se requiere el despliegue de una serie de redes PON, con el fin proveer acceso a la red a un determinado número de usuarios aleatoriamente localizados dentro de esa

región, puede ser tratado como un "grafo de conexión con pesos" (i.e. con conexiones dimensionadas de acuerdo a algún parámetro que permite determinar su "costo" o "peso"). En dicho grafo, las calles de la ciudad y sus intersecciones constituyen los bordes y los nodos del grafo, mismos que pueden ser utilizados como rutas de conexión desde las OLT, hasta las ONU, pasando por sus correspondientes PSC y SSC.

Para describir matemáticamente el problema de optimización que modela la búsqueda del despliegue de redes PON en una ciudad al menor costo posible, es necesario primero definir las siguientes constantes, variables y conjuntos:

- CO: El conjunto de oficinas centrales, $CO=\{CO_C\}$, con $c\in\{1,2,\ldots,C\}$; donde C es el número de CO disponibles.
- N_C: El número de usuarios atendidos por la oficina central c, de manera que $\Sigma N_C = N$, donde N es el número total de usuarios en la región.
- O: El conjunto de OLT, $O=\{O_O\}$, con $o\in\{1,2,\ldots,M\}$, donde M es el número de OLT disponibles.
- U: El conjunto de ONU, $U=\{ONU_n\}$, con $n\in\{1,2,\ldots,N\}$, donde N es el número de ONU requeridas.
- W: El conjunto de longitudes de onda, $W=\{\lambda w\}$, con $w\in\{1, 2,\ldots,L\}$; donde L es el número de longitudes de onda disponibles en el transmisor de una OLT. Específicamente: $L=1$ para *GPON* y *XGPON*, $L=4$ para *NGPON*2 y $L=256$ para *UDWDM PON*.
- L_i: El conjunto de splitters disponibles en un gabinete localizado en el sitio i. También se define en este contexto al parámetro $S_{i,l}$ como el $l^{ésimo}$ splitter, en el gabinete i, cuyo rango de división ("splitting ratio" – *SR*) se denota como $K_{i,l}$ y es igual a 2^r, donde r es número entero positivo.
- B: Es el conjunto de sitios candidatos para colocar un *SSC*.
- V: Es el conjunto de sitios candidatos para colocar un *PSC*.
- n_{max}: Es el máximo número de usuarios por *OLT*.
- ODN_{loss}: La máxima atenuación, en *dB,* permitida en la *ODN*.

Parámetros relacionados con la capacidad de la red PON se definen como sigue:

- BR_{ref} : Bitrate de referencia (para normalización).
- $BR_{US/\lambda}$: La bitrate total para *US* por *OLT* y por cada λ.
- $BR_{DS/\lambda}$: La bitrate total para *DS* por *OLT* y por cada λ.
- BR^{n}_{US} : La bitrate total demandada por la *ONU* $n \in Uc$ en dirección *US*.
- BR^{n}_{DS} : La bitrate total demandada por la *ONU* $n \in Uc$ en dirección *DS*.
- T_{US}: La capacidad total normalizada de bitrate en dirección *US* para una *OLT*, tal que $T_{US} = (L \cdot BR_{US/\lambda}) / BR_{ref}$
- T_{DS}: La capacidad total normalizada de bitrate en dirección *DS* para una OLT, tal que $T_{DS} = (L \cdot BR_{DS/\lambda}) / BR_{ref}$
- γ^{n}_{US}: La demanda de bitrate normalizada de la *ONU* $n \in Uc$, en dirección *DS*, tal que $\gamma^{n}_{US} = BR^{n}_{US} / BR^{n}_{ref}$
- γ_{DS} La demanda de bitrate normalizada de la O*NU* $n \in Uc$, en dirección DS, tal que $\gamma^{n}_{DS} = BR^{n}_{DS} / BR_{ref}$

Además, se empleará un conjunto de parámetros relacionados con los sitios, rutas físicas y costos. Para el efecto se supondrá que en la región bajo estudio ST es un conjunto de calles apropiadas para enrutar cables ópticos y BL es el conjunto de edificios (i.e. cualquier lugar donde haya uno o más usuarios demandando conectividad):

- *I*: El conjunto de nodos en calles y en edificios (intersecciones y vértices), $I = \{i \in \{ST, BL\} / i = 1,2,\ldots,T\}$; donde *T* es el número total de nodos.
- *E*: El conjunto de bordes (enlaces entre nodos) $E = \{e_{i,j} \in E / (i, j) \in I\}$.
- α^{c}_{o}: es una constante binaria que indica (si su valor se 1) que la *OLT* $o \in O$ está ubicada en la oficina central $c \in CO$.
- $d_{i,j}$ La distancia total entre dos nodos $(i,j) \in I$ no necesariamente adyacentes.
- C^{F}_{OF}: Es el costo por unidad de longitud de un cable óptico de alimentación (Feeder OF).
- C^{D}_{OF}: Es el costo por unidad de longitud de un cable óptico de distribución (Distribution OF).

- C_T: El costo por unidad de longitud del zanjado
- C^i_{encl}: El costo de un gabinete óptico en el sitio *i*
- $C_{i,l}$ El costo del $l^{ésimo}$ splitter en el gabinete *i*.
- $C^{rack,\eta}_{OLT}$: El costo de un rack de *OLTs* con capacidad total de hasta η usuarios.
- C^{card}_{OLT}: El costo de un tranceiver *OLTs*.
- $C_{i,l}$ El costo del $l^{ésimo}$ splitter en el gabinete *i*.
- C_{ODF}: El costo de un gabinete de distribución óptico (*ODF*).
- C_{ONU}: El costo de una *ONU*.
- C^{c}_{lbr}: El costo de las labores de instalación y configuración del hardware *PON* en la oficina central *c*.
- α_{OF}: La máxima por unidad de longitud en la *OF.*
- $\alpha_{i,l}$ Atenuación en el $l^{ésimo}$ splitter ubicado en el gabinete *i*.
- α_{ex} La atenuación insertada en la *ODN* por otras fuentes.

Adicionalmente, el modelo de optimización propuesto en este trabajo requiere la definición de las siguientes variables binarias:

$$x_{n,j} = \begin{cases} 1, & \text{si la ONU } n \text{ está conectada al SSC } j \\ 0, & \text{de otra manera} \end{cases}$$

$$x_{j,i} = \begin{cases} 1, & \text{si el SSC } j \text{ está conectada al PSC } i \\ 0, & \text{de otra manera} \end{cases}$$

$$x_{i,o} = \begin{cases} 1, & \text{si un splitter en el PSC } i \text{ está conectado a la OLT } o \\ 0, & \text{de otra manera} \end{cases}$$

$$\alpha_i = \begin{cases} 1, & \text{si el sitio candidato } i \epsilon \{V \cup B\} \text{ está activo} \\ 0, & \text{de otra manera} \end{cases}$$

$$\alpha_o = \begin{cases} 1, & \text{si la OLT } o \text{ está activa (conectada)} \\ 0, & \text{de otra manera} \end{cases}$$

$$S_{i,l} = \begin{cases} 1, & \text{si el } l^{ésimo} \text{ splitter en el sitio } i \text{ está conectado} \\ 0, & \text{de otra manera} \end{cases}$$

$$y_n^{j,l} = \begin{cases} 1, & \text{si la ONU } n \text{ se conecta al el } l^{ésimo} \text{ splitter } i \\ 0, & \text{de otra manera} \end{cases}$$

$$y_{j,l}^{i,p} = \begin{cases} 1, \text{si el } l^{ésimo} \text{ splitter ubicado en el sitio } j \text{ se} \\ \text{conecta con el } p^{ésimo} \text{ splitter del sitio } i \\ 0, \quad \text{de otra manera} \end{cases}$$

$$z_n^o = \begin{cases} 1, & \text{si la ONU } n \text{ se conecta con la OLT } o \\ 0, & \text{de otra manera} \end{cases}$$

B. Formulación del problema de optimización

La función objetivo del problema de optimización apunta a minimizar el costo total de implementación de las múltiples PON que deberán ser desplegadas a partir de varias CO en una gran zona urbana extensa con gran número de usuarios en ella. Una ventaja de optimizar toda la extensa zona en el mismo problema es que se puede variar la distribución de las subregiones servidas por cada CO con el fin de buscar la resolución óptima del problema. En la siguiente sección explicaré cómo se resolvió a través de un enfoque heurístico este problema combinatorio. En media se tendrá N / C usuarios en cada subregión servida por las CO (N es el número total de usuarios y C el número de CO en la región extensa a ser servida).

La función objetivo es la que se muestra en la Ecuación (1).

$$\min \sum_{c \epsilon CO} \Bigg(C_{lbr}^{c} + C_T \Bigg(\sum_{o \epsilon O} \sum_{i \epsilon V} \alpha_o^c x_{i,o} d_{i,o} + \sum_{i \epsilon V} \sum_{j \epsilon B} \alpha_o^c x_{j,i} d_{j,i} \Bigg) + C_{OF}^{f} \sum_{o \epsilon O} \sum_{i \epsilon V} \alpha_o^c x_{i,o} d_{i,o} + \sum_{i \epsilon \{V \cup B\}} C_{OLT}^{card} \alpha_o + C_{OF}^{d} \Bigg(\sum_{i \epsilon V} \sum_{j \epsilon B} x_{j,i} d_{j,i} + \sum_{j \epsilon B} \sum_{n \epsilon U} x_{n,j} d_{n,j} \Bigg) + \sum_{i \epsilon \{V \cup B\}} \sum_{l \epsilon L_i} S_{i,l} C_{i,l} + \sum_{i \epsilon \{V \cup B\}} C_{encl}^{r} \alpha_i + \frac{N}{\eta} \left(C_{OLT}^{rack,\eta} + C_{ODF} \right) + C_{ONU} N \Bigg) \qquad (1)$$

La ecuación (1) está compuesta por una suma global que calcula el costo de implementación para cada subregión servida por las CO. La suma de costos incluye, respectivamente: los costos de labor en la CO y de despliegue de la OF tanto para el zanjado como para el costo de los propios cables, el costo del hardware PON en la CO, el costo de los splitters y gabinetes de calle, y el costo de las ONU.

El conjunto de restricciones que aseguran que el modelo esté apegado a una implementación realista de la red son:

- El número de usuarios conectados a la CO debe ser igual a:

$$N_C = \sum_{n \epsilon U} \sum_{o \epsilon O} \alpha_o^c z_n^o \; ; \quad \forall\, c \in CO \tag{2}$$

- El número de usuarios por OLT debe ser menor o igual a:

$$\sum_{n \epsilon U} z_n^o \leq n_{max} \alpha_o ; \quad \forall\, c \in CO \tag{3}$$

- La demanda de bitrate máxima por OLT no debe ser mayor que su capacidad US y DS:

$$\sum_{n \epsilon U} z_n^o \gamma_{US/DS}^n \leq \Gamma_{US/DS} \alpha_o ; \quad \forall\, o \in O \tag{4}$$

- Una ONU debe estar conectada a un solo splitter que esté ubicado en un determinado sitio j:

$$\sum_{n \epsilon U} x_{n,j} = 1 ; \quad \forall\, n \in U \tag{5}$$

- Un SSC debe estar conectado a un y solo un PSC:

$$\sum_{i \epsilon V} x_{j,i} = \alpha_j ; \quad \forall\, j \in B \tag{6}$$

- Un splitter en un PSC debe llegar a una y solo una OLT:

$$\sum_{o \epsilon O} x_{i,o} S_{i,l} = \alpha_i ; \quad \forall\, i \in V \tag{7}$$

- El número de splitters activos en un sitio i debe ser menor o igual a la capacidad del gabinete:

$$\sum_{n \epsilon U} s_{i,l} \leq \alpha_i r_i; \quad \forall\, i \in \{V \cup B\} \tag{8}$$

- Una ONU se puede conectar a un splitter en el sitio i si existe una conexión física desde la ONU al sitio i.

$$y_n^{i,l} \leq x_{n,i}; \quad \forall\, i \in \{V \cup B\}, \forall\, n \in U, \forall\, l \in L_i \tag{9}$$

- Un splitter l en un SSC ubicado en el sitio j se puede conectar con otro splitter p en un PSC ubicado en el sitio i, si existe una conexión física desde i a j.

$$y_{j,l}^{i,p} \leq x_{j,i}; \quad \forall\, i \in V, \forall\, p \in L_i, j \in B, \forall\, l \in L_j \tag{10}$$

- El número de ONU que se conectan al lesimo splitter activo de un SSC en el sitio j no puede exceder la capacidad del splitter.

$$\sum_{n \epsilon U} y_n^{j,l} \leq K_{j,l} S_{j,l}; \quad \forall\, j \in B, \forall\, l \in L_j \tag{11}$$

- El número de ONU más splitters secundarios que se conectan a un splitter primario, no puede exceder la capacidad del splitter

$$\sum_{n \epsilon U} y_n^{j,p} + \sum_{j \epsilon B} \sum_{l \epsilon L_i} y_{j,l}^{j,p} \leq K_{i,p} S_{i,p}; \quad \forall\, i \in V, \forall\, p \in L_i \tag{12}$$

- La atenuación desde una ONU a su respectiva OLT debe ser menor o igual a la máxima atenuación permitida en la ODN:

$$\begin{aligned} &\alpha_{FO}\left(\sum_{j \epsilon B} x_{n,j} d_{n,j} + \sum_{i \epsilon V} \sum_{j \epsilon B} \sum_{o \epsilon O} x_{n,j} x_{i,o} x_{j,i} d_{j,i} + \sum_{i \epsilon V} \sum_{o \epsilon O} x_{i,o} d_{j,o}\right) \\ &\quad + \sum_{j \epsilon B} \sum_{l \epsilon L_j} y_n^{j,l} S_{j,l} \alpha_{j,l} + \sum_{i \epsilon V} \sum_{p \epsilon L_i} \sum_{j \epsilon B} \sum_{l \epsilon L_j} y_n^{j,l} y_{j,l}^{i,p} S_{i,p} \alpha_{i,p} \\ &\quad + \alpha_{ex} \leq \mathrm{ODN}_{loss} \end{aligned} \tag{13}$$

4. Algoritmo de Solución Desarrollado

A. Función primaria de OTS

El problema descrito por la Ecuación (1), se trata de un problema de optimización del tipo Árbol Mínimo de Steiner, el cual es un bien conocido problema del tipo NP-hard, lo que significa que a medida que no tiene solución exacta o no se puede calcular en un tiempo finito [23]. De modo que para encontrar una solución he desarrollado un algoritmo que hace uso de una serie de heurísticas. El algoritmo trabaja con una función principal (PF) y una serie de funciones secundarias (SF) que son ejecutadas programadamente para que realicen una determinada tarea. A este algoritmo lo he denominado Optimal Topology Search (OTS) [18], mismo que trabaja usando datos de mapas reales provistos por la plataforma de uso abierto OpenStreetMaps™ (OSM). Esto le da la capacidad a OTS de resolver problemas de prácticamente cualquier lugar en el mundo.

La PF carga los datos de OSM (i.e. las coordenadas de estructuras pertenecientes a calles y edificios, incluyendo sus atributos). Dicha información es utilizada tanto para graficar el mapa de la zona donde se realiza el despliegue de la fibra, como para los cálculos que realiza el algoritmo.

En el mapa descargado, cuando se usa el modo simulación (i.e. sin usar datos reales de clientes, su ubicación y demanda de bitrate), PF genera aleatoriamente usuarios en los edificios (dependiendo de su tamaño y tipo de edificio: residencial o corporativo) y les asigna aleatoriamente una demanda de ancho de banda, dependiendo del escenario que se quiere simular (ver Tabla 5). La Figura 3 ilustra una porción de mapa con usuarios.

La generación aleatoria de usuarios se hace en base al área de los edificios (que se puede calcular en base a las coordenadas de su perfil) y al número de pisos de este (información proporcionada directamente por OSM).

OTS tiene la capacidad de resolver el problema para un gran número de usuarios (en el orden de los cientos de miles) todos ellos con heterogéneas demandas de ancho de banda.

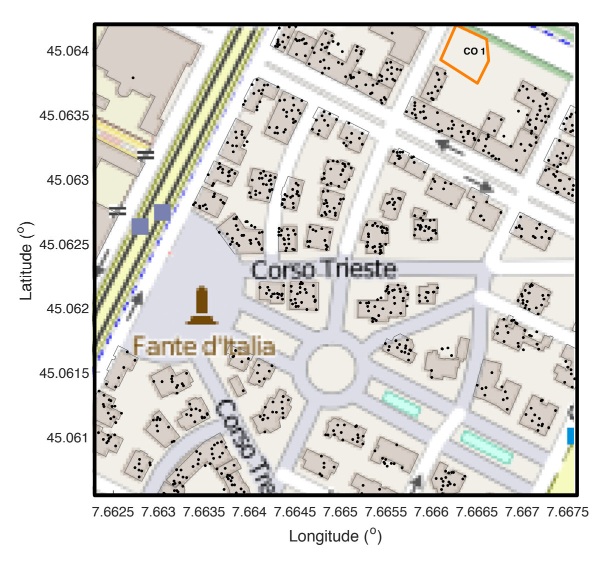


Figura 3. Mapa de ciudad con usuarios en edificios residenciales y corporativos (uno de este tipo se ve en la esquina superior izquierda) generados por OTS

A continuación PF agrupa los usuarios en grandes clúster alrededor de las CO a través de una teselación del tipo Voronoi de la región [24]. Cada clúster significa una subregión servida por su respectiva CO. Inicialmente el centro de cada clúster corresponde al centro de masa de los usuarios, el cual no corresponde (a menos que coincidencialmente coincida) a la ubicación del edificio de la CO. A través de las SF (que son escritas en detalle a continuación) se evalúa el costo del despliegue de múltiples PON usando esa cluterización inicial, y se guarda ese costo. Luego, se realiza una variación de los centros de los clúster (lo que significa un cambio de su tamaño y forma) en dirección a la ubicación geográfica de la CO, y se vuelve a calcular el costo. Este nuevo costo es comparado con el costo anterior, y si resulta mejor, se lo registra como el "menor costo" hasta ahora encontrado y se salva su topología, de otra manera se descarta. Así sucesivamente el algoritmo busca con iteraciones y finura programables el menor costo. El algoritmo converge luego de un determinado número de iteraciones que no se ha encontrado un costo menor (por defecto cinco, pero se puede cambiar ese valor

para dar más confiabilidad, aunque eso aumenta la exigencia del cómputo). Dicho procedimiento constituye una heurística del tipo Tabu Search [25].

En la Figura 4 se ilustra un mapa con tres CO y sus respectivas subregiones iniciales (definidas por las líneas inter-punteadas de color negro) y posteriormente los mapas con clúster modificados, luego de varias iteraciones del algoritmo, donde los centros de masa de los clúster se han movido acercándose a la ubicación de los edificios de las CO.

B. *Funciones secundarias de OTS*

El conjunto de SF, llamadas por PF, son las siguientes:

- *allocate_ssc*: Identifica los edificios que tienen usuarios (residenciales o corporativos) y les asigna un SSC una acometida desde la calle más cercana al edificio (de entre todas las que circundan los mismos). En la Figura 5 se observa una región con varios edificios a los cuales la función allocate ssc les ha asignado un SSC (triángulos rojos) incluyendo su acometida (líneas de color rojo) perpendicular hasta la calle más cercana (las calles son resaltadas por las líneas de color azul en la gráfica).
- *clustr_build*: Divide una subregión de usuarios en clusters, pero no lo hace agrupando usuarios directamente sino a edificios. Se utiliza un enfoque del tipo Shared Nearest Neighbor (SNN) [26], en el cual los edificios son considerados como entidades individuales con atributos numéricos y categóricos. Esta función secundaria de OTS utiliza tres atributos específicamente: el tipo de usuarios en este, el número de usuarios, y el ancho de banda agregado que demandan los usuarios dentro del edificio. Así los usuarios dentro de los edificios terminan agrupados según los parámetros antes mencionados, lo cual es muy útil para dimensionar adecuadamente el hardware PON necesario para atender sus requerimientos de ancho de banda.

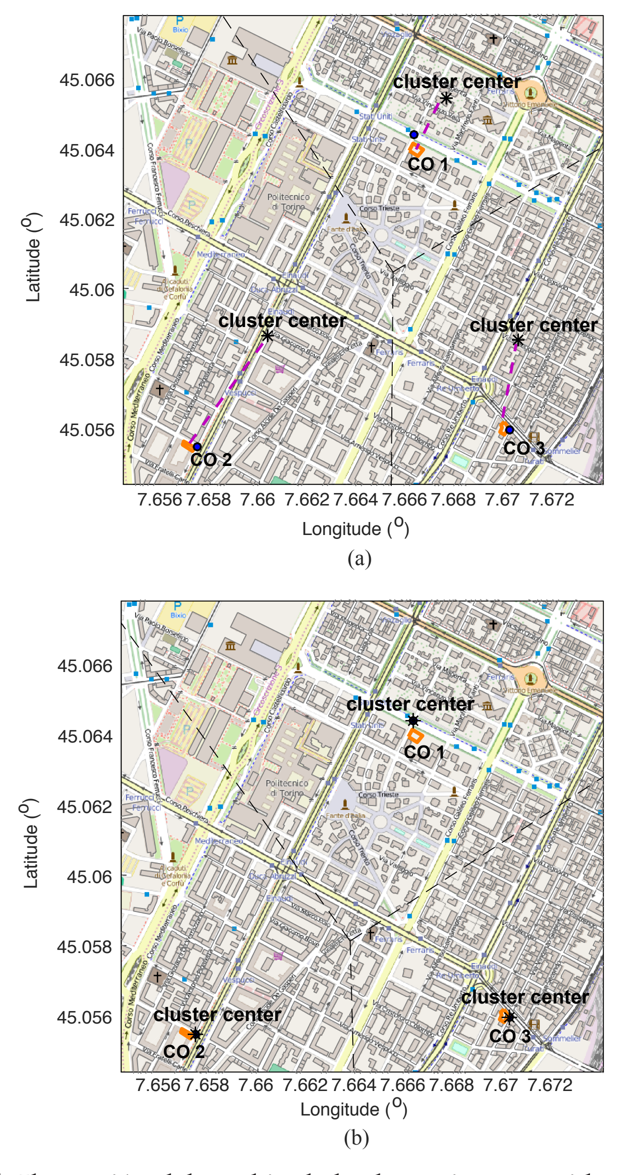


Figura 4. Ilustración del cambio de los las regiones servidas por las CO, en la medida que OTS mueve su centro de masa en cada iteración. (a) Centros de los clúster lejos de los edificios de sus respectivas CO, (b) Centros de los clúster cerca de los edificios de las CO

En la agrupación se toma en cuenta la capacidad de la tecnología PON que se esté planificando desplegar (máximo ancho de banda y número máximo de usuarios que se puede servir con una OLT de la tecnología PON que se emplee). En la Figura 6 se observa la asignación de usuarios y asignación de los mismos a una determinada PON (a una OLT) a través de la agrupación de edificios (cada color corresponde a una PON). Este enfoque es realista dado que en la práctica no es eficiente conectar a la mitad de usuarios de un edificio a una OLT y la otra mitad a otra OLT.

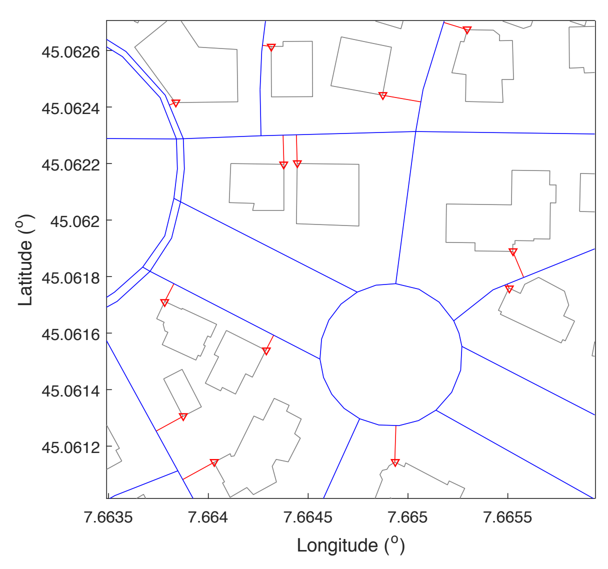


Figura 5. Ilustración de la asignación de SSC a edificios y acometida de OF a la calle más cercana

- *aggregate*: Dimensiona la cantidad y ubicación de los PSC en la ciudad a través de la clusterización de PON (las PON que fueron previamente dimensionadas y asignadas a usuarios por la función anterior). Para el efecto se utiliza una serie de sitios candidatos y nuevamente una teselación de Voronoi (esta vez dentro de la subregión servida por una sola CO). Los centros de masa de los clusters se emplean para la selección inicial del sitio para colocar el PSC, y a través de una búsqueda utilizando un

algoritmo del tipo Random Search, que modifica la cantidad de PSC en cada iteración, se dimensiona de forma óptima los PSC en una subregión. La Figura 7 muestra la agrupación de PON a través de una teselación de Voronoi (líneas inter-punteadas de color negro) y la ubicación óptima de los PSC.

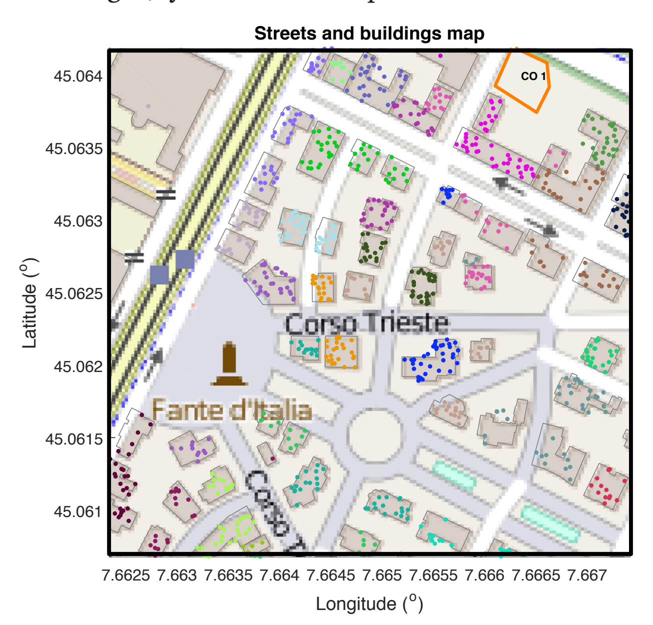


Figura 6. Clusters de usuarios para asignación de hardware PON. Cada color corresponde a un grupo de ONU de usuarios asignados a la misma OLT (cada color corresponde a una PON individual)

- *OF_feeder;OF_distrb;Share*: Estas tres funciones evalúan el zanjado, la compartición de ductos, y las rutas óptimas para el despliegue de los cables ópticos desde las CO hasta los usuarios finales. Para el efecto se utiliza una versión modificada del algoritmo de Dijkstra, el cual busca la iterativamente las rutas más cortas utilizando el criterio no solo de la distancia sino de la recursividad de una ruta (debido a que se debe también optimizar el zanjado y la compartición de ductos). El algoritmo modificado de Dijkstra modifica la matriz de costos en cada iteración disminuyendo el costo de las rutas más recursi-

vas (haciéndolas más favorables en cada iteración). La Figura 8 muestra el trabajo de estas funciones.

- *Evaluate_cost*: Esta función evalúa el costo del despliegue de las PON en la región bajo estudio, incluyendo todas las componentes especificadas en la Eq. (1), utilizando para el efecto costos actualizados y realistas obtenidos directamente de operadores y vendedores. El costo calculado por esta función es enviado a la PF para que lo utilice como criterio de decisión para decidir si la topología actual es menor (o no) que la que actualmente está registrada como la topología de menor costo.

5. Resultados

A. Función primaria de OTS

OTS encuentra una solución óptima de menor costo para el despliegue de múltiples PON en zonas urbanas extensas con muchos usuarios y con variadas demandas de ancho de banda. Dado que se trata de una solución basada en heurísticas no es seguro que se trata de la solución óptima absoluta, pero sí provee de una solución óptima cuya confiabilidad depende del número de iteraciones, de la precisció de la búsqueda y del tipo de barrido que se programe (muchos pasos de corta distancia, pero sin embargo ello implica más tiempo computacional consumido en la búsqueda de la solución).

OTS devuelve dos tipos de resultados: numéricos y gráficos. Los resultados numéricos incluyen tablas detalladas de la distribución de PON entre edificios y usuarios, como: a qué SSC y PSC están conectadas, la ubicación y dimensión de los SSC y PSC y la longitud y tipo de cables ópticos. Se presenta además información detallada de los costos, tanto de las componentes de cada subregión, así como de toda la zona evaluada.

Con el fin de presentar los resultados de un específico escenario de solución se ha escogido una región del centro de varias ciudades, escogiendo zonas que tengan potencialmente alrededor de 100000 usuarios.

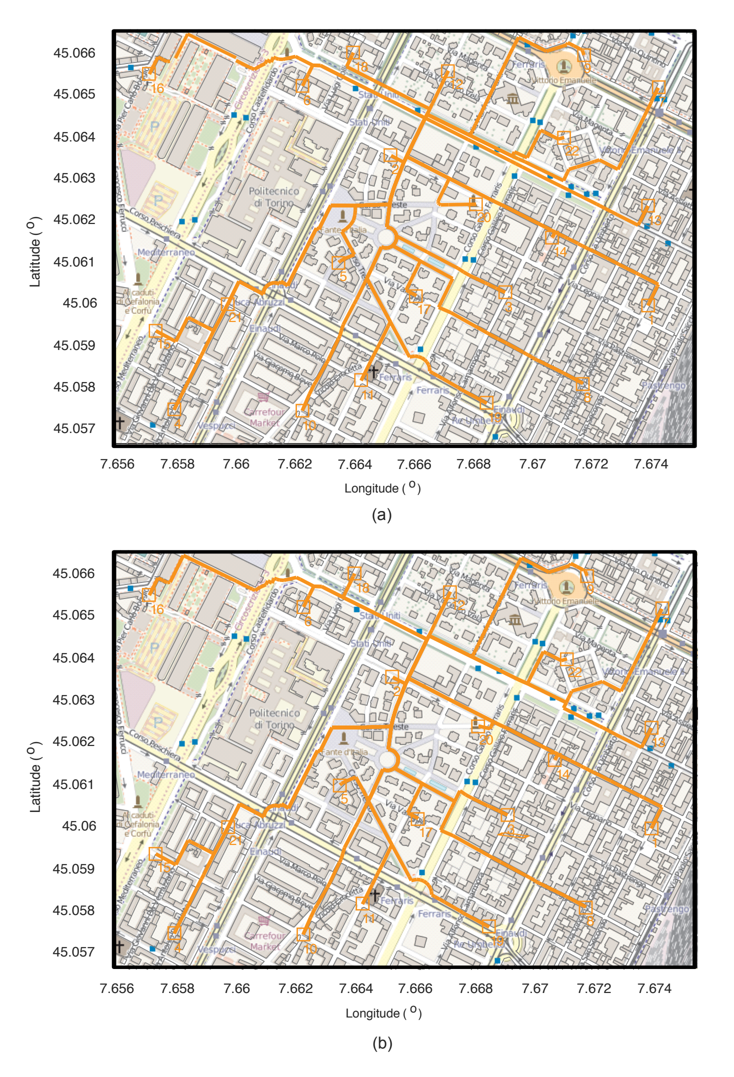


Figura 8. Función *OF_feeder*: (a) Rutas desde la CO hasta los PSC generadas por la primera iteración de la función (usando una versión modificada de Dijkstra) (b) Rutas óptimas determinadas una vez que el algoritmo converge

En dichas zonas se corrió OTS para encontrar la solución de menor costo para el despliegue de todas las tecnologías PON especificadas en la Tabla 4.

En la Figura 9 se muestra la solución encontrada por OTS en una porción de la ciudad de Turín (Italia). Se observa en las gráficas la posición de los PSC, los SSC, así como las rutas de los cables ópticos (de feeder y distribución). En la tabla VI se detalla los resultados de costos de implementación para las tecnologías PON consideradas y para cada uno de los seis bitrate especificados en la Tabla V.

Algoritmo 1: Optimal Topology Search (OTS)

```
Data:    Data = load_data (City,Users,CO)
Result: Optimal_Topology=OTS(Data)
begin
   for  i ∈ {Heuristic_modifier_counter} do
      Data_i = i^th_heuristic_variation (Data)
      for  c ∈ CO  do
           SSC = allocate_ssc (Data_i)
           PON_hardware = cluster_buildings (SSC, Data_i)
           PSC = aggregate (PON)
           OF_cables = find_paths (SSC, PSC, Data_i)
           ODN = {OF_cables, PSC, SSC}
           Trenching = share (ODN)
           Topology_i = {ODN, PON_hardware }
      end
      C_i = evaluate_cost (Topology_i)
      if  C_i < C_opt  then
            Optimal_topology = Topology_i
            C_opt = C_i
      end
   end
end
```

Tabla VI: Costo del despliegue para $2 \cdot 10^5$ usuarios

Escenario	COSTO (millones de USD)			
	GPON	XGPON	NGPON2	UDWDM PON
#1	51.4	82.2	113.3	157.1
#2	61.0	84.2	113.3	157.1
#3	106.6	93.3	113.3	157.1
#4	178.7	127.9	113.3	157.1
#5	394.6	207.8	146.8	159.9
#6	-	250.7	168.3	164.5



Figura 9. Resultado gráfico entregado por OTS: Una porción de una subregión de la ciudad de Turín, en la que se observa la ubicación óptima de los PSC, los SSC asignados a cada edificio, y los cables ópticos enrutados en las calles

Nótese que una región con cerca de 100000 usuarios es realmente mucho más grande que la región del mapa de ejemplo mostrada en la Figura 9. En tal región, realmente se tienen varias subregiones servidas cada una de ellas por una CO diferente, con aproximadamente 20000 usuarios en cada subregión. En la Figura 10 se muestra un ejemplo de una zona urbana con tres CO (los edificios de cada subregión están pintados de un distinto color). Se trata solo de un ejemplo con fines ilustrativos ya que en la práctica las subregiones, con cerca de 20000 usuarios, son mucho más grandes.

En la Figura 11 se puede apreciar una comparación de los costos de implementación para las cuatro tecnologías PON consideradas en este estudio y para los seis escenarios de demanda de ancho de banda (bitrate) detallados en la Tabla V. Es interesante verificar que los resultados del escenario 1 (usuarios residenciales demandando entre 10 a 50 Mb/s de ancho de banda), que corresponde a un escenario de demanda real en el presente, da como resultado valores bien conocidos en la actualidad: GPON es la tecnología más redituable, ya que su costo de despliegue corresponde a aproximadamente $500 por cada usuario (un costo absolutamente real hoy

en día para el despliegue de GPON), mientras que XGPON es más costosa (no se justifica la inversión aún) y mucho más costosas en este escenario son la NGPON2 y la UDWDM PON.

En el escenario 2 (usuarios residenciales demandando entre 50 a 100 Mb/s), escenario cercano al actual (real en algunas ciudades del primer mundo y casi real en la mayoría de las urbes), GPON se convierte en una solución un poco más costosa, sin embargo, las otras tres tecnologías siguen siendo más caras de implementar (nótese de todos modos que la brecha de costos se ha reducido pues no ha cambiado el precio de implementación de las otras tecnologías, solo ha subido el costo de GPON).

El escenario 3 (usuarios residenciales demandando entre 100 a 400 Mb/s), que es un escenario a corto plazo (pocos años en el futuro), implementar la tecnología GPON deja de ser la mejor opción ya que se torna más costosa que la XGPON, NGPON2 es no mucho más costosa que GPON y XGPON, lo cual es muy interesante dada la considerablemente mayor escalabilidad de NGPON2 frente a las otras dos.

Para el escenario 4 (usuarios residenciales demandando entre 100 Mb/s a 1 Gb/s), el cual es por supuesto un escenario a largo plazo (sobre todo por la consideración de tener usuarios que demanden en sus casas velocidades cercanas o iguales a 1 Gb/s), NGPON2 es claramente la mejor opción frente a las otras tecnologías. Se trata de un punto de inflexión que hace pensar en la conveniencia de implementar esta tecnología, ya estandarizada, incluso desde el escenario anterior.

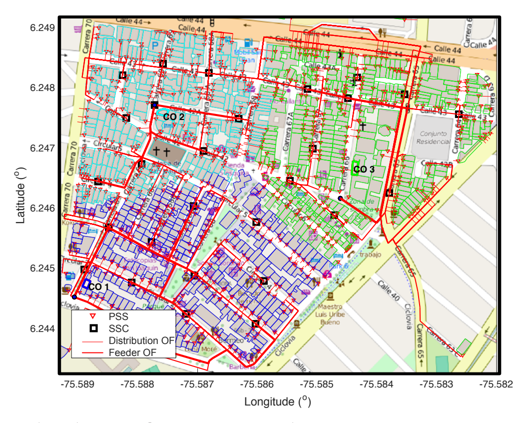


Figura 10. Resultado gráfico entregado por OTS: Una zona de la ciudad de Medellín, Colombia, con tres subregiones

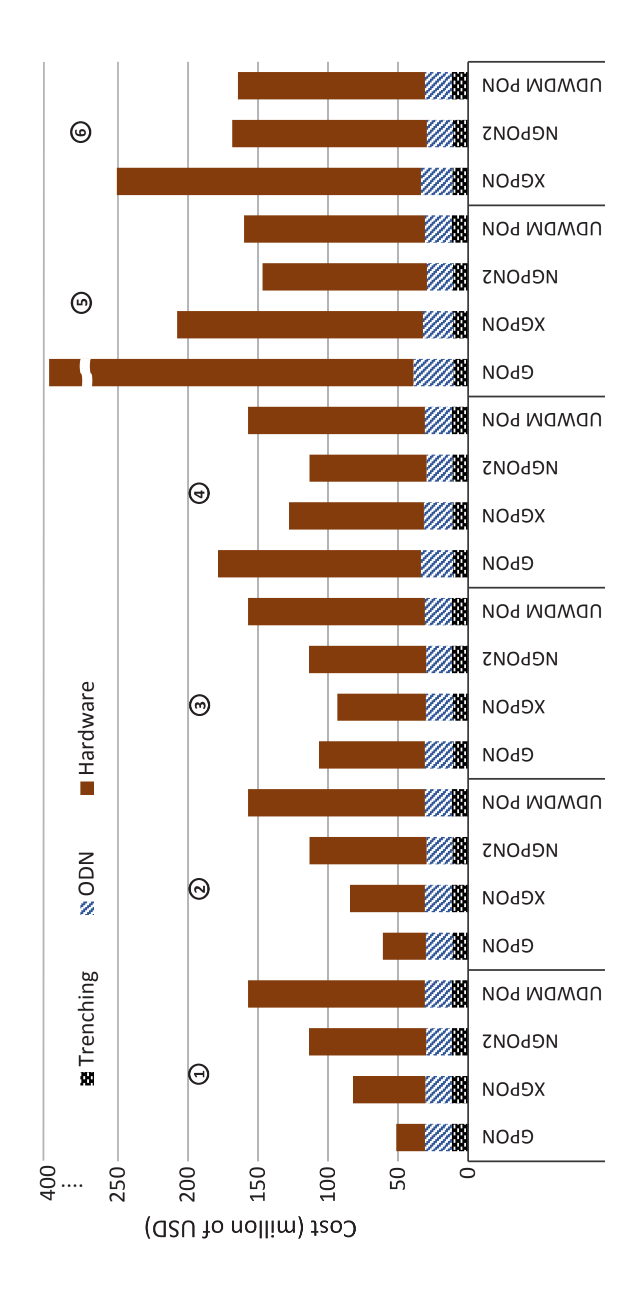


Figura 11. Costo total global y detallado para tres grandes componentes: zanjado (trenching), ODN y Hardware, del despliegue óptimo de GPON, XGPON, NGPON2 y UDWDM PON en los seis escenarios de bitrate detallados en la Tabla V en una zona urbana densa con aproximadamente 100000 usuarios

Los escenarios 5 y 6 son escenarios a muy largo plazo (el 6 más demandante que el 5 dado que pretende que los usuarios requieran al menos 1 G/s, mientras que el 5 supone una demanda desde los 500 Mb/s). En el escenario 6 el aumento de costo de GPON es tan alto que incluso no es práctico incluir a esa tecnología en la gráfica para dicho escenario. Es poco predecible en cuántos años en el futuro los usuarios podrían demandar estos enormes anchos de banda, pero desde este análisis queda claro que solo entonces tendría sentido pensar en implementar una tecnología tan costosa como lo sería la UDWDM PON (aunque es la que potencialmente puede entregar más ancho de banda a los usuarios en un futuro aparentemente lejano).

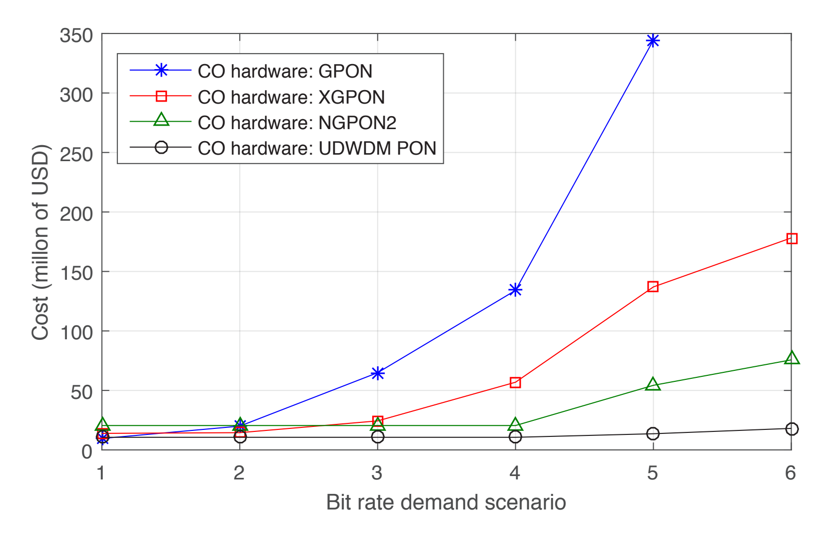


Figura 12. Costo total para aproximadamente 1000 usuarios del hardware PON en la CO para GPON, XGPON, NGPON2 y UDWDM PON en los seis escenarios de demanda de bitrate

La Figura 12 permite apreciar algo que en la figura anterior es también evidente: el aumento en el costo de implementación de una tecnología, con el aumento del ancho de banda, se debe principalmente al aumento en el costo del hardware PON, sobre todo en la CO. Esto se debe a que cuando una demanda de ancho de banda, por parte de

los usuarios, comienza a sobrepasar los límites de una tecnología, esta comienza a convertirse en una solución costosa y poco redituable. Es importante ver en la figura que las tecnologías que casi no cambian su costo de implementación (aunque inicialmente mayor) son NGPON2 y UDWDM PON (esta última casi no cambia de costo de hardware incluso en el escenario 6, lo que significa que es la solución más económicamente escalable de todas). Las otras dos, GPON y XGPON, incrementan su costo exponencialmente en los seis escenarios considerados.

6. Conclusiones

El presente trabajo de investigación evidencia la utilidad, versatilidad y confiabilidad de OTS para el análisis tecno-económico del despliegue óptimo de redes PON en zonas urbanas extensas con heterogéneas y cambiantes demandas de ancho de banda por parte de los usuarios.

OTS puede ser utilizado para el análisis del despliegue óptimo no solo de PON, sino de todo tipo de red de información, como enlaces punto a punto o fronthauling para redes inalámbricas de nueva generación, lo que constituye potencialmente nuevos trabajos de investigación utilizando esta herramienta.

Se pueden llegar a conclusiones muy relevantes gracias al empleo de OTS en un análisis tecno-económico, que soporten la toma de decisiones de inversión, tales como los escenarios en los cuales el despliegue de una determinada tecnología sería o no redituable y el dimensionamiento de la escalabilidad de una solución tecnológica frente al cambio de las demandas de los usuarios.

Referencias

[1] E. Wong, “Next-Generation Broadband Access Networks and Technologies,” *J. Light. Technol.*, vol. 30, no. 4, pp. 597–608, Feb. 2012.
[2] Q. Guo and A. V. Tran, “20-Gb/s Single-Feeder WDM-PON Using Partial-Response Maximum Likelihood Equalizer,” *IEEE Photonics Technol. Lett.*, vol. 23, no. 23, pp. 1802–1804, Dec. 2011.
[3] G. Arévalo and D. Cárdenas, “Advantages of the use of VCSEL over RSOA for uplink transmission on WDM-PON networks,” in *2012 IEEE Latin-America Conference on Communications, LATINCOM 2012 - Conference Proceedings*, 2012.

[4] K. Y. Cho, S. Member, U. H. Hong, M. Suzuki, and Y. C. Chung, "103-Gb / s Long-Reach WDM PON Implemented by Using Directly Modulated RSOAs," *IEEE Photonics Technol. Lett.*, vol. 24, no. 3, pp. 2011–2013, 2012.

[5] J. Prat, "Technologies for a Cost Effective UDWDM-PON," in *Procceedings of OFC*, 2015, pp. 1–3.

[6] R. Gu, X. Liu, H. Li, and L. Bai, "Evolutional algorithm based cascade long reach Passive Optical Networks planning," *China Commun.*, vol. 10, no. 4, pp. 59–69, Apr. 2013.

[7] L. Gouveia, M. J. Lopes, and A. de Sousa, "Single PON network design with unconstrained splitting stages," *Eur. J. Oper. Res.*, vol. 240, pp. 361–371, Jul. 2014.

[8] J. Z. J. Zhang and N. Ansari, "Minimizing the Arrayed Waveguide Grating Cost and the Optical Cable Cost in Deploying WDM Passive Optical Networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 1, no. 5, pp. 352–365, Oct. 2009.

[9] R. Chowdhury and B. Jaumard, "A cross layer optimization scheme for WDM PON network design and dimensioning," *2012 IEEE Int. Conf. Commun.*, pp. 3110–3115, Jun. 2012.

[10] J. Segarra, V. Sales, V. Polo, and J. Prat, "Dimensioning OLT Architectures for UDWDM-PONs Employing Coherent Transceivers," in *ICTON 2015: 17th International Conference on Transparent Optical Networks*, 2015, no. Ga 318515, pp. 2–7.

[11] D. Nesset and S. Member, "NG-PON2 Technology and Standards," *J. Ligthwave Technol.*, vol. 33, no. 5, pp. 1136–1143, 2015.

[12] D. L. Truong, P. T. Do, and A. T. Pham, "Optimization of Survivable Mesh Long-Reach Hybrid WDM-TDM PONs," *J. Opt. Commun. Networking, OSA Publ.*, vol. 6, no. 1, pp. 62–76, 2014.

[13] O. Kipouridis, C. M. Machuca, A. Autenrieth, and K. Grobe, "Street-aware infrastructure planning tool for Next Generation Optical Access networks," in *6th International Conference on Optical Network Design and Modeling (ONDM)*, 2012, pp. 1–6.

[14] S. Pal, C. Zukowski, A. Nag, D. B. Payne, and M. Ruffini, "Cable Length Minimisation in Long-Reach-PON Planning for Sparsely Populated Areas," in *International Conference on Optical Network Design and Modeling*, 2014, pp. 19–22.

[15] G. V. Arévalo, R. C. Hincapié, and J. E. Sierra, "Optimization model for UDWDM-PON deployment based on physical restrictions and asymmetric users' clustering," in *Proc. SPIE 9626, Optical Systems Design 2015: Optical Design and Engineering VI*, 2015, vol. 9626, pp. 1–11.

[16] A. Eira, J. Pedro, and J. Pires, "Optimized Design of Multistage Passive Optical Networks," *J. Opt. Commun. Networking, OSA Publ.*, vol. 4, no. 5, pp. 402–411, 2012.

[17] G. V. Arévalo, J. E. Sierra, R. C. Hincapié, and R. Gaudino, "A novel algorithm for PON optimal deployment over real city maps and large number of users," in *IET Conference Publications*, 2016, vol. 2016, no. CP704.
[18] G. V. Arévalo, R. C. Hincapié, and R. Gaudino, "Optimization of multiple PON deployment costs and comparison between GPON, XGPON, NGPON2 and UDWDM PON," *Opt. Switch. Netw.*, vol. 25, 2017.
[19] ITU-T, "G.984.1 Gigabit-capable passive optical networks (GPON): General characteristics," 2008.
[20] R. Murano, "Optical Component Technology Options for NGPON2 Systems," *OFC2014*, no. c, pp. 3–5, 2014.
[21] H. Rohde, E. Gottwald, A. Teixeira, J. D. Reis, A. Shahpari, K. Pulverer, and J. S. Wey, "Coherent Ultra Dense WDM Technology for Next Generation Optical Metro and Access Networks," *J. Light. Technol.*, vol. 32, no. 10, pp. 2041–2052, 2014.
[22] H. Rohde, S. Member, E. Gottwald, E. Weis, P. Wagner, Y. Babenko, D. Fritzsche, and H. Chaouch, "Trials of a Coherent UDWDM PON Over Field-Deployed Fiber: Real-Time LTE Backhauling, Legacy and 100G Coexistence," *J. Ligthwave Technol.*, vol. 33, no. 8, pp. 1644–1649, 2015.
[23] M. R. Garey, R. L. Graham, and D. S. Johnson, "The Complexity of Computing Steiner Minimal Trees," *SIAM J. Appl. Math.*, vol. 32, no. 4, pp. 835–859, 1977.
[24] M. De Berg, M. Van Kreveld, and M. Overmars, "Computational geometry: algorithms and applications," *Springer-Verlag, New York,* vol. 3ra.ed., no. ISBN: 978-3-540-77973-5, pp. 191–218, 2008.
[25] F. Glover, M. Laguna, and R. Marti, "Principles of tabu search," *Approx. Algorithms Metaheuristics*, vol. 23, pp. 1–12, 2007.
[26] L. X. L. Xia, J. S.-Y. J. Sheng-Yi, and S. X.-K. S. Xiao-Ke, "A Novel Fast Clustering Algorithm," in *International Conference on Artificial Intelligence and Computational Intelligence*, 2009, vol. 4, pp. 284–288.

Sobre los autores



Germán V. Arévalo recibió su título de ingeniería en Electrónica por parte de la Escuela Politécnica Nacional, Quito – Ecuador, en el 2003. Recibió su título de Máster en Ciencias en Comunicaciones Ópticas por parte del Politécnico di Torino, Turín – Italia, en el 2004 y su título de PhD, por parte de la Universidad Pontificia Bolivariana de Medellín – Colombia, en el 2016. Actualmente es profesor titular principal y director de la Carrera de Ingeniería en Telecomunicaciones y del programa de Maestría de Investigación en Telemática, en la Universidad Politécnica Salesiana (UPS) del Ecuador. Es a su vez Coordinador del Grupo de Investigación en Electrónica y Telemática, GIETEC, de la UPS. Su investigación se ha desarrollado alrededor de las redes ópticas de nueva generación, a nivel experimental y teórico.



Roberto Gaudino recibió su título de ingeniería en Electrónica, con Summa Cum Laude, por parte del Politécnico di Torino (POLITO), Turín – Italia, en 1993 y su PhD en el POLITO, en 1998. Es actualmente profesor asociado del POLITO. Ha dirigido en el pasado proyectos colaborativos con TIM Italia, que han significado contribuciones al nuevo estándar de redes PON, NGPON2 y hacia el estándar futuro de PON (conocido hoy como HS-PON). Ha sido coordinador de varios proyectos europeos, como el POF-ALL y el POF-PLUS, enfocados en redes ópticas plásticas de bajo costo. Participó en importantes proyectos como el ALPHA y el E-PhotonOne (donde actuó como coordinador del WP en "Comunicaciones de Corto Alcance"). Además, se ha desempeñado como director del programa de Ingeniería en Telecomunicaciones del POLITO y actualmente es el coordinador científico del PhotoNext Center, que se enfoca en comunicaciones ópticas de ultra-alta velocidad sensores ópticos y componentes para aplicaciones industriales. PhotoNext constituye uno de los 11 más grandes centros departamentales del POLITO, cuyo financiamiento comenzó en el 2017.

CAPÍTULO VII

Clasificación automática de nódulos mamográficos basada en fusión de información visual multi-vista

Automatic classification of mammographic nodules based on multi-view information fusion

Fabián Narváez*
Universidad Politécnica Salesiana Ecuador
*email: fnarvaeze@ups.edu.ec

Resumen

Una correcta evaluación e interpretación mamográfica demanda un alto nivel de experticia del radiólogo observador, depende directamente de un adecuado análisis visual de los hallazgos y la correlación de las características radiológicas extraídas de diferentes proyecciones mamográficas. En este artículo se presenta un esquema de clasificación automática de nódulos contenidos en Regiones de Interés (RoIs), extraídos de dos diferentes proyecciones mamográficas (Medio Lateral Oblicua y Creaneo-Caudal) obtenidas de la misma glándula mamaria, mediante un estrategia de fusión de información ipsilateral. Una vez que, el radiólogo especialista selecciona una Región de Interés en las dos proyecciones mamográficas, estas son caracterizadas mediante descomposiciones multi-resolución y multi-escala, para lo cual, cada RoI es proyectada sobre dos espacios diferentes definidos por los momentos de Zernike y la transformada Curvelet, respectivamente. Así esta información heterogénea es optimamente fusionada mediante una estrategia de aprendizaje de múltiples kernels (Multiple Kernel Learning) construida por un entrenamiento con máquinas de soporte vectorial (SVM). El desempeño de la estrategia propuesta para clasificar nódulos malignos y benignos, fue evaluado respecto a un esquema de clasificación basado en el análisis de la RoI correspondiente a una única proyección, para lo cual se utilizaron un conjunto de 980 RoI extraídos de 490 casos de estudio de la base de datos de mamografía (DDSM) y 216 RoI extraídos de 108 casos de estudio de la base de datos INBreast. Los resulta-

dos obtenidos reportan una sensibilidad del 98.3% y especificidad del 94.5% frente a una sensibilidad del 96.2% y especificidad del 93.1% obtenidos cuando se utilizaron características multi-resolución en una única proyección. Estos resultados sugieren que la estrategia propuesta puede ser útil en escenarios clínicos y puede contribuir a la capacitación de nuevos radiólogos como segundo lector.

Palabras clave

Diagnóstico asistido por computador, multiple kernel learning, transformada curvelet, momentos de zernike.

Abstract

Correct mammography assessment and interpretation demands great expertise of radiologist observer and depends directly on a suitable visual analysis of mammographic findings and their correlation with radiographic features extracted from different mammographic views. In this paper, an automatic classification scheme for mammographic nodules contained on Regions of Interest (ROIs) is presented, which is based on an information fusion approach by using RoIs extracted from two different mammographic views of the same breast, a Mediolateral Oblique (MLO) view and a craniocaudal (CC) view, respectively. Once the expert radiologist selects a RoI from the two mammographic projections, those are characterized by using a multiresolution and multiscale decomposition approaches. For doing so, each RoI is projected into two different spaces defined by Zernike moments and Curvelet transform, respectively. Finally, this extracted heterogeneous information is optimally fused by using a Multiple Kernel Learning strategy based on Support vector machine scheme. The performance of the herein proposed strategy, for classifying benign and malignant nodules, was evaluated respect to the classical mammographic analysis based on only mammographic view, for which a set of 980 ROIs extracted from 490 cases and other set of 216 RoI extracted from 108 cases, which were extracted from DDSM and INBreast databases, respectively. The obtained results reported a sensitivity of 98.3% and a specificity of 94.5% respect to 96.2% and 93.1% of sensibility and specificity, respectively, and obtained by the analysis based on an only mammographic view. These results suggest that the herein proposed strategy could be useful in real clinic scenarios and could be contributing to the training of new radiologists.

Keywords

Mammography, computer aided diagnosis, multiple kernel learning, curvelet transform, zernike moments.

1. Introducción

La mamografía es aún considerada la mejor técnica costo-efectiva para detectar anormalidades relacionadas a la presencia de cán-

cer de mama en etapas tempranas de desarrollo [1, 2]. El tamizaje a través de mamografías ha demostrado ser efectivo reduciendo la mortalidad por el cáncer de mama con porcentajes que varían desde el 30 al 70% [3]. Sin embargo, la sensibilidad diagnóstica durante el tamizaje mamográfico es ampliamente afectada por: la calidad de las imágenes mamográficas, la similitud de los tejidos normales frente a los anormales visualizados en la mamografía y la experiencia del radiólogo observador, lo que produce una alta variabilidad inter e intra-observador durante la interpretación de los hallazgos mamográficos. Estudios previos han reportado que la sensibilidad del tamizaje mamográfico varía entre el 85% y 90%, mientras que otros estudios han encontrado un importante porcentaje de presencia de falsos positivos durante los procesos diagnósticos, entre el 10% y 30% [4, 5, 6].

Durante las últimas décadas han sido propuestas diferentes soluciones con el propósito de incrementar la exactitud durante la interpretación mamográfica, su sensibilidad y especificidad, así como, reducir las tasas de biopsias innecesarias causadas por ciertas imprecisiones diagnósticas, entre estas soluciones importantes, el Colegio Americano de Radiología (American College of Radiology, ACR) desarrolló un sistema para el reporte diagnóstico mediante una terminología estandarizada para los informes de hallazgos anormales en imágenes mamarias (BI-RADS) mediante una semiología radiológica, para de esta manera asignar una categoría entre 0 y 6, de acuerdo a sus características radiológicas específicas. Estas categorías apoyan a los médicos y determinan la probabilidad de malignidad para considerar ciertas recomendaciones finales para el manejo y control de los hallazgos [7].

En el caso particular de las masas mamarias, de acuerdo al BI-RADS se describen por su forma, margen y densidad. En la quinta edición de BI-RADS, la forma se describe como ovalada, redonda e irregular y el margen como circunscrito, oscurecido, microlobulado, espiculado o indefinido. Finalmente, la densidad, a la cual se le considera como la atenuación de la masa con respecto al tejido glandular, se puede describir como Alta densidad, Igualdad de densidad, Baja densidad o Grasa. La figura 1 ilustra el espectro de características visuales que determinan el nivel de probabilidad de

malignidad de acuerdo a la formas de masa y márgenes descritos en el sistema BI-RADS.

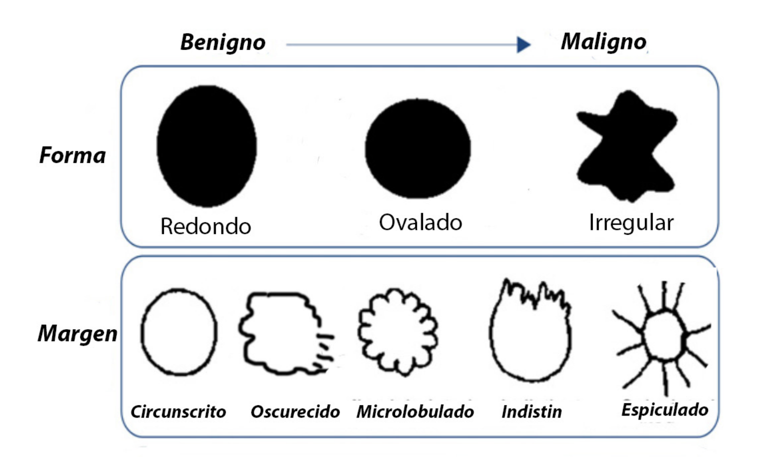


Figura 1. Diferentes tipo de formas y márgenes de masa definidos en el BI-RADS. El grado de malignidad depende de estas características morfológicas

Como ilustración, una masa con forma redonda u ovalada, margen circunscrito y baja densidad tiene una alta probabilidad de ser benigna, mientras que una masa con forma irregular, margen espiculado y alta densidad es sospechosa de cáncer [8]. Este valor también define el manejo particular del paciente, es decir, los valores bajos implican que el paciente es mantenido con un adecuado control periódico mientras que los valores altos generalmente conducen a la posibilidad de biopsia y estudio patológico adicional. Sin embargo, la interpretación mamográfica sigue siendo una tarea subjetiva y difícil para el radiólogo.

Por otro lado, se han desarrollado tecnologías de Diagnóstico Asistido por Computador (CAD) como herramientas computacionales para asistir al radiólogo durante la toma de decisiones y son usadas como segundos lectores. En la actualidad, estos sistemas CAD son considerados como un potencial soporte al diagnóstico y tamizaje mamográfico, los mismos que han demostrado su utilidad durante la identificación de agrupamientos de micro-calcificaciones [9, 10], reportando una sensibilidad que varía entre el 75 al 90%,

pero su desempeño para soportar el diagnóstico de masas (nódulos) han sido reportados muy bajos, con una sensibilidad que varía entre el 54 al 72% [11, 12]. La baja sensibilidad de los CAD comerciales y sus altas tasas de presencia de falsos positivos han reducido la confianza del radiólogo en estos sistemas [12]. La mayoría de los CAD están diseñados para producir una opinión sobre la presencia o no de cáncer, pero su análisis se basa únicamente mediante el procesamiento de la información de una sola proyección mamográfica, pudiendo ser ésta, la proyección Medio Lateral Oblicua (MLO) o la proyección Cráneo-Cudal (CC) de la misma mama, mientras que el protocolo clínico en los escenarios reales, el flujo de trabajo del radiólogo durante la interpretación de las lesiones mamográficas se basa en el análisis del hallazgo obtenido en las dos proyecciones, MLO y CC, para así determinar las características relevantes patológicas de la lesión y emitir un adecuado diagnóstico. Esto debido a la alta variabilidad de las características radiológicas que pueden presentar los hallazgos anormales vistos en diferentes proyecciones. Debido a que los sistemas CAD solo analizan una proyección, el uso clínico de este tipo de herramientas sigue siendo muy limitado, ya que la evidencia no es concluyente y, a veces, incluso contradictoria. Varios investigadores han descubierto que los sistemas CAD pueden aumentar la detección de las etapas iniciales de la enfermedad [13], mientras que otros informaron que, en un entorno de detección bien establecido, el uso de dichos sistemas no ayudó significativamente a mejorar la detección del cáncer [12].

Por lo tanto, varios estudios han intentado mejorar el rendimiento de los sistemas CAD, especialmente en el caso de análisis de masas [9, 14]. En la literatura, se han reportado varios algoritmos automáticos para segmentación y caracterización de masas dirigidos a tareas de clasificación como benignas o malignas [15, 16, 14, 17], reportando resultados con valores de AUC, área bajo la curva ROC, de aproximadamente 0.8 al combinar diferentes tipos de características de bajo nivel [17]. Típicamente, los descriptores de bajo nivel aplicados en mamografías son construidos mediante la extracción de las características morfológicas y de textura de las lesiones mamarias, previamente seleccionadas por el radiólogo. Entre la gran cantidad de enfoques reportados en la literatura se encuentran algunos descripto-

res basados en los momentos de Zernike, los cuales se han utilizado para caracterizar formas y márgenes de masa previamente segmentadas y contenidas en regiones de interés [18, 19]. Además, los momentos de Zernike se han usado ampliamente en diferentes tipos de aplicaciones como un descriptor general de forma [20, 21, 22, 23] en masas bien definidas [18]. Del mismo modo, se han utilizado una gran variedad de técnicas de procesamiento de imágenes basadas en información estadística, estructurales y espectrales para analizar mamogramas [24], incluido el análisis por matrices de coocurrencia [25], dimensiones fractales [26, 27], wavelets [28, 29], curvelets [30] y contourlets [31, 32], mostrando un rendimiento muy variable.

Todos estos enfoques abordan el problema de clasificar una masa como maligna o benigna. Sin embargo, en la mayoría de los sistemas CAD se ignoran las diferencias visuales que pueden presentar la misma masa vista en varias proyecciones por lo que han limitado su análisis a las características visuales únicas en una única proyección, las cuales son evaluadas de forma independiente. Por lo tanto, la información entre las múltiples proyecciones o vistas complementarias de la misma mama, desafortunadamente no es considerada. La figura 2 ilustra la diferencia entre las características visuales de una misma masa observada en diferentes proyecciones.

Este hecho presenta como resultado inexactitudes y una alta incidencia de tasas de falsos positivos (PF). Recientemente, se ha propuesto el uso de información de las dos proyecciones mamográficas para mejorar el rendimiento del CAD respecto a una sola proyección. Hasta ahora, los sistemas CAD tradicionales se utilizan para identificar regiones mamográficas sospechosas en cada proyección por separado, de las que se extraen las características morfológicas y de textura. Estas características se utilizan luego para determinar la correspondencia entre el par de regiones extraídas de las dos proyecciones de la mama, ya sea utilizando diferentes clasificadores discriminativos, como el análisis discriminante lineal (LDA) y el k-vecino más cercano (K-NN). Estos enfoques demostraron una capacidad para discriminar entre correspondencias verdaderas y falsas, reduciendo los falsos positivos reportados por el uso de un sistema CAD de una única proyección. Aunque el uso de información de múltiples proyecciones ha demostrado ser útil para mejorar la detec-

ción de masas mamográficas, sin embargo, estos enfoques dependen en gran medida de una segmentación precisa de la lesión, un desafío difícil durante el análisis de masas con límites no definidos.

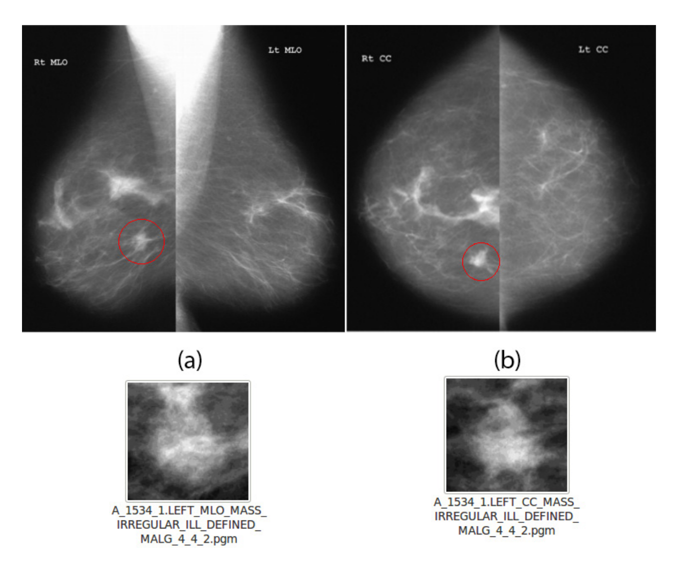


Figura 2. Caso de estudio mamográfico: a) Ilustra una RoI con una masa extraída de la proyección MLO de la mama izquierda b) Ilustra una RoI con la misma masa extraída de la proyección CC de la mama izquierda

Este artículo presenta un método que clasifica automáticamente las masas como benignas o malignas contenidas en mamografías multi-vista. En lugar de llevar a cabo una segmentación completa de la lesión, la descripción automática se basa en una caracterización multi-resolución y multi-escala de la región de interés (RoI) extraída de al menos dos proyecciones MLO y CC, así cada RoI es proyectada sobre dos espacios de transformación diferentes definidos, el primero por los momentos de Zernike y el segundo por la transformada Curvelet, respectivamente. Finalmente, esta información heterogénea es fusionada mediante una estrategia de aprendizaje de múltiples kernels (Multiple Kernel Learning) construida por un entrenamiento con máquinas de soporte vectorial (SVM). En trabajos anteriores se ha mostrado el potencial de las características

de multi-resolución como descriptores de masas en una proyección mamográfica [33, 34, 35] y en este artículo se amplía estos trabajos generalizando la representación y aprendiendo de las características relevantes extraídas de cada proyección para finalmente correlacionar dicha información, logrando así construir un único descriptor extraído de las dos proyecciones.

Las principales contribuciones de este trabajo se resumen a continuación:

1. Una nueva caracterización de masas mamográficas tomadas de dos proyecciones diferentes y que fusiona dos bases complementarias de descomposición multi-resolución, el análisis piramidal de los momentos de Zernike, una base ortogonal compleja y la transformada Curvelet, una descomposición multi-escala adecuada para capturar mezclas complejas de líneas curvas.
2. Un método que integra varias proyecciones del mismo hallazgo mamográfico para clasificar su nivel de severidad (maligno/benigno) acorde al protocolo clínico real radiológico.

El resto de este artículo está organizado de la siguiente manera: en la siguiente sección se resume una revisión de la literatura de trabajos previos desarrollados para clasificar masas según un análisis ipsilateral. En la sección 3, se describe el método propuesto. La sección 4 presenta los resultados experimentales, y la última sección está dedicada a discusiones y conclusiones.

2. Materiales y métodos

En la figura 3 se ilustra el esquema de trabajo del método propuesto. Inicialmente, el radiólogo experto selecciona, de forma manual, una masa, la misma que está contenida en las dos vistas mamográficas (MLO y CC) de la misma mama, como Regiones de Interés, entonces la estrategia comienza aplicando una etapa de pre-procesamiento para cada una de las Regiones de Interés seleccionadas, una vez que se han mejorado algunos detalles visuales, la información correspondiente a la forma y textura de la masa, es extraída usando los momentos de Zernike y la transformada Curvelet. Esta información es combinada usando una estrategia de aprendizaje por

múltiples kernels (Multiple Kernels Learning) asignando pesos según la relevancia de las características visuales en tareas de clasificación, un procedimiento que se logra mediante un entrenamiento basado en máquinas de soporte vectorial (SVM). Una vez que las regiones son caracterizadas, estas son clasificadas como benigna o maligna.

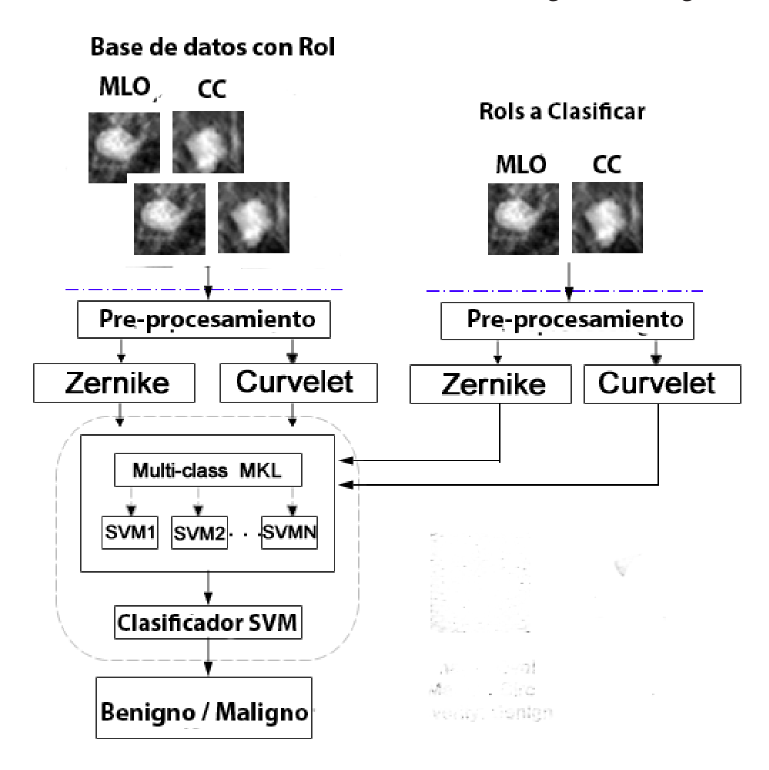


Figura 3. Esquema general de la estrategia propuesta

A. *Preprocesamiento de mamografías*

Inicialmente, las mamografías usadas en este trabajo fueron normalizadas mediante una estrategia de conversión de escala de grises hasta los 8-bits de profundidad [36]. Para este efecto, únicamente se seleccionaron mamografías que contenían algunas masas con un diámetro estimado hasta de 2 cm, las mismas que corresponden a una región cuadrada (RoI) de 512×512 pixeles, las

cuales son más probables a desarrollar cáncer en tapas tempranas y mucho más difíciles de diagnosticar [37]. Fundamentalmente, se ha demostrado que los radiólogos integran la información visual percibida a varias escalas y diferentes frecuencias [38]. En el caso particular de las masas, el análisis mamográfico integra información de los bordes y patrones espaciales locales en diferentes escalas y a diferentes orientaciones. Inspirado en estos estudios, en el presente trabajo se ha explorado dos representaciones complementarias que describen las lesiones mamográficas de manera global (descriptor de forma) y de manera local (descriptor de textura). Para el primer caso, se propone un análisis piramidal de los momentos de Zernike con el propósito de capturar información esencial correspondiente a la forma de la masa y una descomposición mediante Curvelets a diferentes escalas y orientaciones para capturar información de diferentes lineas curvas y patrones de textura, pero para cada descriptor fue necesario un pre-procesamiento diferente, los mismos que se describirán a continuación:.

B. *Descriptor de forma de masas: análisis por los momentos de Zernike*

La forma de las masas ha sido reportada entre los signos radiológicos más importantes [39, 6, 40] y los momentos de Zernike se han usado clásicamente como buenos descriptores para caracterizar formas [20, 22]. Los momentos de Zernike se derivan de un conjunto de polinomios ortogonales complejos, proyectados dentro del círculo unitario [41, 42], los mismos que son expresados en coordenadas polares y calculados desde un conjunto de polinomios de Zernike, tal como:

$$V_{p,q}(r,\theta) = R_{p,q}(r)e^{jq\theta}, r \in [-1,1] \qquad (1)$$

donde, p representa el orden del polinomio y q sus repeticiones (frecuencias). Así, el orden p es un interno no negativo y la repetición q es un entero positivo o negativa, de tal manera que satisface la condición $p-|q|$ es un número par y $|q|\leq p$. Los polinomios radiales con valores reales están dados por:

$$R_{p,q}(r) = \sum_{k=q}^{p} B_{p|q|k} r^k e^{jq\theta} \qquad (2)$$

donde, $r=\sqrt{x^2+y^2}$ es la magnitud del vector $\theta = tan^{-1}(\frac{y}{x})$ y su ángulo, y sus coeficientes $B_{p|q|k}$ están dados por:

$$B_{p|q|k} = \frac{(-1)^{\frac{p-k}{2}}(\frac{p+k}{2})!}{(\frac{p-k}{2})(\frac{k+q}{2})!\,(\frac{k-q}{2})!} \qquad (3)$$

Estas funciones ortogonales, permiten descomponer una imagen $f(x,y)$ como una suma ponderada de los polinomios de Zernike, tal como: $\sum_{p=1}^{\infty} \sum_{q=-p}^{p} Z_{p,q} V_{p,q}$, donde $Z_{p,q}$ son los momentos de Zernike de orden p, con q repeticiones. Esta relación expresada en términos discretos, en donde una imagen digital de tamaño $N \times N$ es proyectada dentro del espacio de Zernike contenido por todas las funciones polinomios ortogonales, se presenta como:

$$Z_{pq} = \frac{p+1}{\lambda} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} V_{pq}^{*}(x,y) f(x,y) \qquad (4)$$

Donde, $\lambda=\delta A/\pi$ es una constante de normalización y δA es el área de la imagen cuando es proyectada sobre el circulo unitario de los polinomios de Zernike.

Como un primer paso, se aplicó el método propuesto en [35], para resaltar las características de forma de la masa, el cual consistió en cuantificar el rango dinámico del histograma de nivel de grises original en 12 bins dentro del intervalo de [0,255]. Para tal efecto, se asigna una RoI particular al círculo unitario haciendo coincidir el centroide de la RoI y el círculo unitario. De esta manera, se calcularon los momentos de Zernike de 45^{th} hasta el 50^{th} orden, siendo estos valores los suficientes para reconstruir la forma original con una relación-señal a ruido pico, PSNR (Peak Signal-to-Noise Ratio),

mayor a 40:dB. Del mismo modo, este orden seleccionado evita cualquier inestabilidad numérica producida por la presencia de altas frecuencias de la imagen, tal como se ha demostrado en [35], siendo esta una limitación de los polinomios de Zernike producidos por la discretización de estas bases continuas. En este trabajo los errores numéricos y geométricos producidos por el uso de esta estrategia se corrigieron como se muestra en [42]. Finalmente, este principio de extracción de características se aplicó para las dos regiones desde las vistas MLO y CC, para finalmente construir un vector de características general mediante la concatenación del histograma normalizado de acuerdo a las frecuencias (repeticiones) para cada uno de los cinco órdenes (45 a 50) de los momentos de Zernike, obtenidos por cada RoI en las diferentes vistas.

Por otro lado, los efectos producidos por la invarianza a la traslación y escalado de la imagen se evitan mediante la aplicación de una técnica de normalización como se describe en [43], al centrar y escalar el marco de referencia (x,y) a $(\frac{x}{a}+\bar{x},\frac{y}{a}+\bar{y})$, siendo $(x,\bar{y})$ el centroide $f(x,y)$ y $a=\sqrt{\beta/m_{00}}$ el factor de escala, con β el número de pixeles en la masa obtenida al binarizar el RoI y m_{00} el momento geométrico de orden cero.

Clásicamente, las representaciones de formas tomadas desde los primeros cinco órdenes de los polinomios de Zernike han sido usados como una aproximación a las aberraciones ópticas producidas por el ojo humano [19, 18], ignorando los coeficientes de orden alto ya que, dependiendo del tamaño de la imagen, un cálculo simple para los primeros 40 momentos pueden tomar una hora [42]. Sin embargo, con la estrategia propuesta en este trabajo, el costo computacional se reduce drásticamente en la fase de pre-procesamiento, los momentos hasta el orden cincuenta pueden calcularse en menos de 5s. Además, este subconjunto de momentos también se seleccionó para obtener una representación de las características globales de masas (ver figura 4).

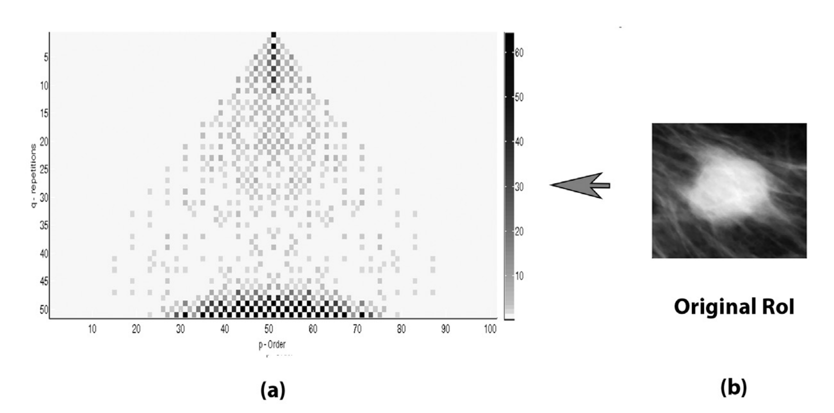


Figura 4. Descomposición piramidal mediante los momentos de Zernike (frecuencia vs. orden): a) Ilustra los coeficientes de Zernike hasta el orden 50 de la imagen original. b) Ilustra una RoI con una masa extraída de la proyección CC

C. *Caracterización por la transformada Curvelet*

La transformada Curvelet realiza una descomposición multiescala [44] utilizando un conjunto de bases estadísticamente dependientes. En donde cada RoI de las proyecciones MLO y CC, son descompuestas para capturar su información de acuerdo a diferentes escalas, de igual manera, cada una de estas bases Curvelet captura su orientación asociada. Esta transformación, en particular, limita cada escala mediante la ley de escalamiento anisotrópico parabólico [44], asegurando que la forma de masa y los patrones de textura se conserven a diferentes escalas. Varias investigaciones han demostrado que Curvelet es un descriptor de textura adecuado en diferentes bases de datos públicas con imágenes de textura [45], mientras que otras han demostrado que Curvelets supera las transformadas como wavelets o Gabor en tareas de clasificación [32, 46]. En el dominio de la frecuencia, un Curvelet se construye como el producto de dos ventanas: una angular y las coronas radiales diádicas. La ventana angular representa el análisis direccional y la ventana diádica radial es un filtro pasa-banda con frecuencias de corte que siguen la ley de escalado anisotrópico parabólico $widh \approx length^2$ [44]. Debido a esta propiedad anisotrópica, los Curvelets se adaptan mucho mejor a

las curvas a diferentes escalas que otras transformaciones y, por lo tanto, resultan muy resistentes al ruido en las imágenes. La figura 5 ilustra la descomposición mediante la transformada Curvelet de una RoI.

El pre-procesamiento para este caso particular de extracción de características, busca preservar los patrones principales de textura y definir un nivel comparable de intensidad, para ello, un filtro basado en la mediana fue aplicado, suavizando la RoI de cada proyección (MLO y CC), mientras que con un ajuste adaptativo del histograma [47] se lograron resaltar los patrones de textura principales y conserva los bordes importantes, como se ilustra en la figura 6.

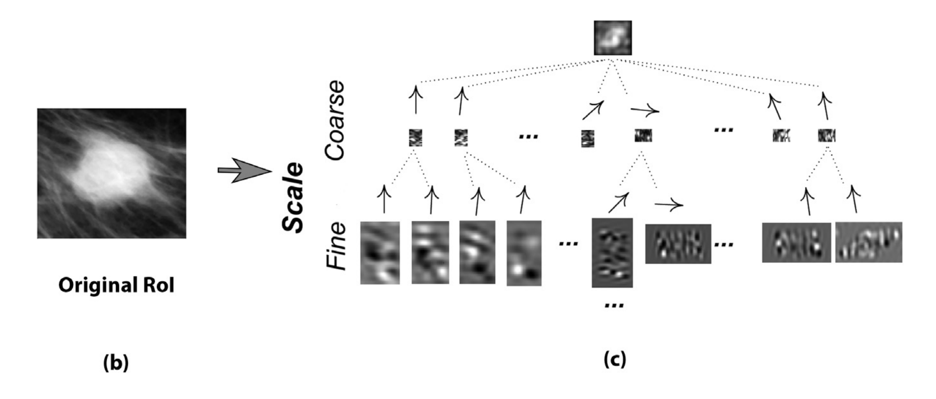


Figura 5. Descomposición multi-resolución mediante la transformada Curvelet de una RoI a diferentes escalas y orientaciones

Luego de que las dos regiones de interés RoI de las dos proyecciones fueron descompuestas en cuatro escalas con 32 orientaciones, para un total de 128 sub-bandas. El descriptor basado en curvelets se construyó concatenando versiones normalizadas de los coeficientes curvelets extraídos de las diferentes subbandas, es decir, 32 orientaciones en las cuatro escalas por la RoI de la proyección MLO y la región de la proyección CC. Para este caso particular y con el propósito de definir la similitud para este tipo de descripción, se ha implementado una medida de similitud basada en la Intersección de Histogramas. La transformada Curvelet fue implementada mediante

la transformada rápida de Fourier desigualmente espaciada, con una ventana rectangular inclinada a lo largo de la dirección principal de cada curvelet, como es presentado en [44]

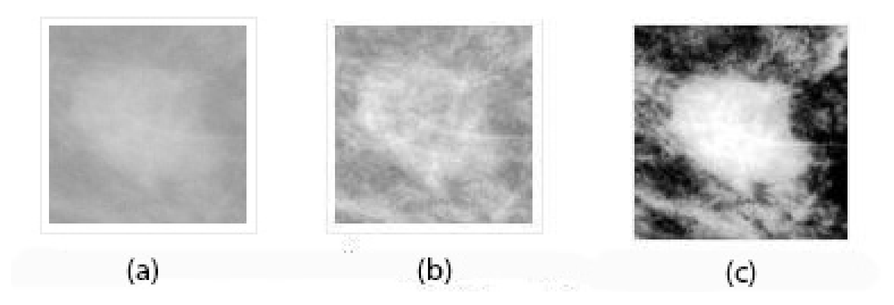


Figura 6. Etapa de procesamiento mediante un filtro de mediana y ajuste adaptativo del histograma. a) Ilustra una RoI original con una masa, b) Ilustra una RoI procesada con un filtro de mediana, c) Ilustra la RoI mejorada por ajuste adaptativo del histograma

D. *Fusión de descriptores multiresolución*

Una vez que se realizaron los análisis de Zernike y Curvelet para cada RoI (MLO y CC), se aplica una estrategia de fusión de información con el propósito de encontrar una forma óptima de combinar la información obtenida de estas dos fuentes de representación. Específicamente, la fusión propuesta intenta encontrar valores óptimos para los parámetros que maximizan la capacidad discriminativa para las clases maligna y benigna de una masa. Una estrategia común es ponderar los espacios de características (descriptores Zernike y Curvelets) y luego sumarlos para calcular una caracterización óptima, tal como [48]:

$$\hat{S}(x) = \sum_{\phi} \omega_{\phi} S^{\phi}(x) \qquad (5)$$

El problema con esta estrategia es que la información importante de las características individuales puede perderse cuando se com-

binan de manera lineal los espacios de características. Por lo tanto, la fusión propuesta en este trabajo usa una estrategia diferente que mantiene toda la información de los dos espacios de características y solo los fusiona cuando se necesita tomar una decisión. Específicamente, los espacios de Zernike y Curvelet se utilizan como entrada para una función discriminante, $DW(x)$, que indica a qué espacio pertenece un determinado valor de x para una clase en particular:

$$D_W(x) = \left\langle W, \left(\Psi_1\left(S^1(x)\right), \ldots, \Psi_\Phi\left(S^{|\Phi|}(x)\right)\right)\right\rangle \qquad (6)$$

Donde, $W \in R^{|\Phi| \times M \times N}$ es el vector de parámetros que indica la importancia relativa del vector de características extraído de cada espacio $S^\varphi(x)$ y $\Psi_{1,\ldots,\phi}$ $:[S^\phi]^{M \times N} \rightarrow F$ es la función que mapea cada descriptor $S^\varphi(x)$ para un espacio de características F. El valor de $D_W(x)$ es esperado a satisfacer $D(x) \geq 0$ sí x corresponde a una específica clase y $D_W(x) < 0$ sí x no es parte de esa clase. Esta formulación tiene la ventaja de mejorar la flexibilidad permitiendo que el modelo tenga en cuenta interacciones no-lineales potencialmente complejas entre los valores de un espacio único original correspondiente a características en diferentes escalas.

Considerando que solo dos espacios de características diferentes deben ser óptimamente fusionados, dos funciones Ψ, específicamente la función Ψ_1 y Ψ_2 son encargadas de mapear los descriptores a un nuevo espacio de características, donde los patrones complejos no-lineales se vuelven lineales. Esta estrategia es bien conocida como *kernel trick* y es usada con métodos de kernels, donde el mapeo Ψ_1 y Ψ_2 es implícitamente inducida por una función kernel. Así, una función kernel $k{:}X \times X \rightarrow R$ asociado a un mapeo $\Psi_{1,\ldots,\phi}{:}X \rightarrow F$ tal que$\forall$ x,y ϵ X,k(x,y)$\leq \Psi_1$ (x),Ψ_1 (y)$>F$, i.e., k calcula el producto punto en F. De esta manera, un Kernel puede ser visto como una función que mide la similitud entre dos objetos desde un espacio de características. En este trabajo, el espacio de características de entrada puede ser el espacio de las Curvelets y los momentos de Zernike que son fusionados, entonces la función kernel mide la similitud en esos espacios particulares.

Esto es formulado como problema de maximizar y minimizar el siguiente problema de discriminación:

$$\min_{W} C \sum_{i=1}^{N} max(0,1 - y_i D_W(x_i)) + \|W\|_2^2 \qquad (7)$$

donde, x_i representa una RoI de entrenamiento, $y_i \in \{-1,1\}$ representa la clase correspondiente, *N* es el número de muestras usadas en el entrenamiento, y *C* controla la regularización del modelo. El primer término de la función objetivo en la ecuación 7 es una función de pérdida que penaliza la clasificación incorrecta de las muestras de entrenamiento y el segundo término es el regularizador del parámetro *W*. La regularización de *W* está asociada con la búsqueda de un máximo margen del clasificador de la misma manera como funciona para la clasificación mediante máquinas de soporte vectorial [49]. En este trabajo, el parámetro C fue de 1000 después de una búsqueda exhaustiva del valor óptimo bajo una estrategia de entrenamiento mediante una validación cruzada con 10 sub-grupos de datos, conocida como “10-fold cross validation”.

Entonces, considerando las funciones Ψ_1 y Ψ_2, en este trabajo se usó la intersección de histogramas como una función kernel, la cual es definida como:

$$\Psi_1 = k_{hi}(s_p, s_q) = \sum_{i} \sum_{j} min(s_p(i,j), s_q(i,j))$$

lo cual requiere que los momentos de Zermike sean normalizados entre valores de 0 a 1. Así, el modelo propuesto fusiona los diferentes espacios combinando sus respectivos kernels, en lugar de directamente sumar los mismo espacios. El problema de optimización en la ecuación 7 es por lo tanto tratado como un problema de aprendizaje basado en múltiples kernels, conocido como Multiple Kernel Learning (MKL), con el cual se busca un buen hiper-plano discriminante en el espacio fusionado. Varias formulaciones basadas en MKL han sido desarrolladas últimamente, la estrategia usada en este trabajo se basa en la propuesta por [50], Aprendizaje Generalizado basado en Multiples Kernel (GMKL).

E. *Clasificación de masas*

Esta etapa de clasificación consiste en establecer sí una nueva masa seleccionada por el radiólogo, en las dos vistas (MLO y CC), corresponde a una masa maligna o benigna de acuerdo a los criterios establecidos por el sistema BI-RADS. Una vez que las RoIs se asignan a los dos espacios diferentes como se describió hasta ahora, las diferentes representaciones alimentan un clasificador binario SVM (Support Vector Machine) que utilizan sus kernels pre-calculados basados en la intersección del histograma como medida de distancia. Este clasificador para discriminar entre las dos clases, la clase benigna y maligna, asume como valores positivos y negativos para cada clase respectivamente, lo cual es determinado por el signo de la siguiente función discriminante [51]:

$$f_i = \frac{\boldsymbol{w}_i . x + \mathbf{b}_i}{||\mathbf{w}_i||}, i = 1,2,\ldots,n \qquad (8)$$

donde w_i y b_i son la normal y el peso hacia el hiper-plano relacionado a el i^{th} clasificador. Por lo tanto, dada las dos RoIs x, el conjunto de clase $c_i = \{benign, malign\}$. es asignado de acuerdo al $sign:(f_i(x))$

F. *Medidas de Evaluación*

En tareas de clasificación, el clasificador debe asignar una etiqueta (clase) de una colección de posibilidades finita y mutuamente exclusivas y, por lo tanto, cada instancia estará asociada a una sola clase. Esto generalmente se encuentra en la literatura como el método de relevancia binaria [51] por lo que se han utilizado diferentes medidas de desempeño. En este trabajo, el desempeño es evaluado en términos de la exactitud (*accuracy*), *sensibilidad*, *especificidad* y el área bajo la curva (AUC) de la curva RoC (Receiver Operating Characteristic) que es una representación gráfica de la sensibilidad frente a la especificidad para un sistema clasificador binario de acuerdo el umbral de discriminación. Estas medidas son definidas como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Sensibilidad} = \text{TPR} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Especificidad} = \frac{TN}{TN + FN} \quad (11)$$

donde *TP*, *TN*, *FP* y *FN*, denotan la cantidad de verdaderos positivos, falsos negativos, falsos positivos y falsos negativos, respectivamente, obtenidos de dicha tarea de clasificación. En consecuencia, se entrenó y evaluó el esquema de clasificación SVM, para el cual se determinaron los parámetros de clasificación durante la etapa de entrenamiento mediante una validación cruzada con 10 particiones de datos (10-fold cross validation).

G. *Descripción de las bases de datos*

La estrategia propuesta es evaluada utilizando diferentes niveles de calidad de imágenes mamográficas, una versión digitalizada de mamografías de placa (screen-film) y mamografías digitales de campo completo. Para esto, se usan dos bases de datos diferentes de dominio público, la Base de datos digital para la detección de mamografías (DDSM)[1] y la base de datos INBreast[2], de las cuales fueron extraídos dos conjuntos de RoIs de las vistas MLO y CC, respectivamente. Una breve descripción para cada conjunto de datos se muestra debajo.

- Base de datos DDSM: la base de datos digitalizada para tamizaje de mamografía (DDSM) [52] es una base de datos con imágenes mamográficas ampliamente usada en esta área de investigación como punto de referencia de evaluación [53, 54, 55]. Esta base de datos de libre acceso está constituida por imágenes digitalizadas de películas mamográficas con su infor-

1 http://marathon.csee.usf.edu/Mammography/Database.html

2 http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database

mación técnica y diagnóstico clínico correspondiente. Toda la base de datos DDSM contiene un total de 2620 casos, cada uno incluye cuatro imágenes obtenidas de las vistas Cranio-Caudal (CC) y Mediolateral-Oblicua (MLO), así como una descripción específica de acuerdo al sistema BI-RADS, y que fueron anotadas por radiólogos expertos según BI-RADS cuarta edición. En esta investigación, se eligió un subconjunto de mamografías, las mimas que cumplieron los siguientes criterios de inclusión: 1) similar calidad de imagen, 2) número equilibrado del tipo de masas (benignas/malignas), 3) solo se evaluó una sola lesión, es decir, este estudio excluyó masas con calcificaciones, distorsión arquitectural o casos con la presencia de clips o marcas de metal en el área de masa. Por lo tanto, fueron seleccionadas un subconjunto de mamografías digitalizadas con un escáner láser Lumisys a 50µm o un escáner Howtek a una resolución de 43.5µm píxeles y con un rango dinámico de intensidades de = 4096 niveles de tono de grises. Luego, un proceso de extracción de la RoI fue llevado a cabo. Para lo cual se consideró la información relacionada a la ubicación y tamaño de la masa de cada caso, permitiendo recortar manualmente subimágenes cuadradas considerando como centro las masas. Debido a que las anotaciones proporcionadas en la base de datos DDSM incluyen el límite de la segmentación de los hallazgos, en este trabajo el tamaño de las RoIs se definió como un cuadrado de nxn pixeles. Específicamente, las RoI se recortaron según el cuadro delimitador de las lesiones y se redimensionaron a nxn píxeles conservando la relación de aspecto cuando el ancho o la altura del cuadro delimitador son mayores que n; en otros casos, la lesión se centra sin escalar y se preservan los tejidos de fondo. En consecuencia, un conjunto de 980 regiones de interés fueron extraídas de esta base de datos DDSM incluyendo vistas CC y MLO, distribuidas como 490 regiones de interés con masa de la vista MLO y 490 regiones de interés de la vista CC de la misma mama.

- Base de datos INBreast: una evaluación adicional fue llevada a cabo usando los casos almacenados en la base de datos INBreast [56]. Esta base de datos contiene mamografías en

formato DICOM y fueron adquiridas del Hospital de São João, Breast Centre, Porto, Portugal. Esta colección de casos incluye masas, calcificaciones, asimetrías bilaterales y distorsión arquitectural. En este trabajo, se seleccionaron un total de 108 cases (216 imágenes) que constituyen esta base de datos, de los cuales 108 casos son documentados ipsilateralmente (4 imágenes por caso) y 25 son de pacientes con mastectomía (2 imágenes en estos casos). En consecuencia, un conjunto de 108 masas fueron extraídas por cada vista MLO y CC. La localización y el tamaño de las masas, asociadas a cada caso, fueron tomados de los contornos dibujados por los expertos y proporcionados como información en formato XML.

3. Resultados

El método propuesto fue evaluado dos veces. Primero, para la tarea de clasificar RoIs de entrada tomadas de la base de datos DDSM y determinar sí corresponde a la clase benigna o maligna. En segundo lugar, la evaluación fue llevada a cabo con las RoIs tomadas de la base de datos INBreast para la misma tarea de clasificación. En ambos casos, se evaluó el desempeño del método de fusión de información de las dos proyecciones (MLO y CC) respecto a la clasificación obtenida por cada vista. El proceso de evaluación experimental fue implementado en MATLAB R17, ejecutado en una PC Linux con 2 Intel Quad Core i7 a 3.07 GHz y 24 GB de RAM. Todos los experimentos fueron implementados y basados en el código fuente de Generalized Multiple Kernel Learning (GMKL), así como, la implementación de la estrategia Multiple Kernel Learning [50], mientras que el clasificador SVM fue implementado usando la librería de código abierto LIBSVM [57], con matrices de distancia como kernels pre-calculados.

A. *Base de Datos DDSM*

La primera evaluación del método consistió en comparar el poder discriminante del esquema de fusión de información obtenida de las dos proyecciones MLO y CC frente a la caracterización de la misma masa únicamente vista en cada proyección por separado, para tal efecto, una misma masa fue caracterizada mediante los

espacios de Zernike y Curvelets en la proyección MLO y CC. Esta evaluación permitió determinar el rendimiento del esquema de fusión propuesto frente a imágenes con calidad variable, como es el caso de las imágenes de la base de datos DDSM. En esta evaluación se usaron 980 RoIs extraídas de la base de datos DDSM (490 MLO y 490 CC). Por lo tanto, en esta evaluación se comparó el desempeño obtenido con la proyección medio lateral oblicua (MLO), proyección cráneo caudal (CC) y la estrategia de fusión de las dos proyecciones (MLO-CC). Las dos espacios de características se usaron como entradas para un clasificador SVM binario clásico con matrices de kernel pre-calculadas. Para el proceso de entrenamiento se llevó a cabo una partición de 70-30 muestras, para entrenamiento se seleccionó el 70%, mientras que el 30% restante se usó para la validación, un proceso que fue repetido 10 veces. Los resultados obtenidos durante esta evaluación se presentan en la Tabla 1.

Tabla 1. Rendimiento promedio según un esquema de validación de 70–30% para la tarea de clasificar los RoI como benignas y malignas para las base de datos DDSM

	DDSM		
	MLO-CC	**MLO**	**CC**
Sensibilidad	96.3	85.3	82.2
Especificidad	93.1	83.2	80.4
AUC	0.95	0.83	0.79
Accuracy	97.8	86.1	80.3

En general, la fusión de información correspondiente a las dos proyecciones (MLO-CC) supera al desempeño logrado por cada una de las proyecciones por separado MLO y CC, respectivamente, alcanzando una sensibilidad del 96.3% frente 85.3% y 82.2% correspondiente a la proyección MLO y CC, mientras las especificidades alcanzada de 93.1 fue superior a 83.2% y 80.4% alcanzados por cada proyección por separado. Además, las métricas *Accuracy* y *AUC* muestran resultados notables, con 97.8 y 0.95 para esa base de datos. Finalmente, estos resultados pueden ser observados mediante las curvas ROC, los cuales se ilustran en la figura 7

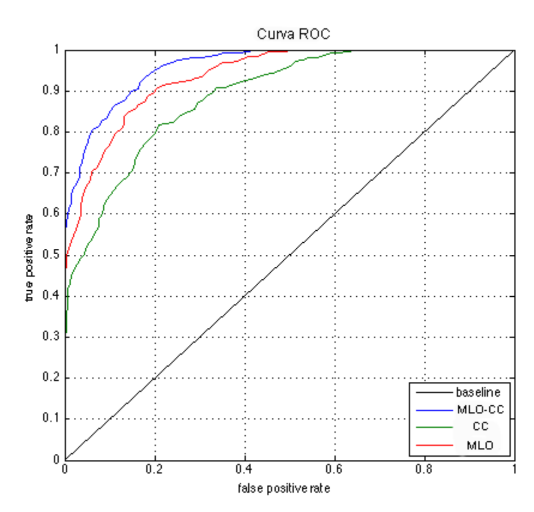


Figura 7. Curva ROC correspondiente al desempeño de la estrategia de fusión propuesta (MLO-CC) respecto a la clasificación usando información de una única vista MLO y CC, respectivamente

B. *Base de Datos INBreast*

En esta evaluación, se midió el desempeño de la estrategia propuesta en imágenes digitales DICOM con la misma calidad. De igual manera, esta evaluación consistió en comparar el poder discriminante del esquema de fusión de información obtenida de las dos proyecciones MLO y CC frente a la caracterización de la misma masa únicamente vista en cada proyección por separado. En esta evaluación se usaron 216 RoIs extraídas de la base de datos INBreast (108 MLO y 108 CC). De igual manera, se comparó el desempeño obtenido con la proyección medio lateral oblicua (MLO), proyección cráneo caudal (CC) y la estrategia de fusión de las dos proyecciones (MLO-CC). Las dos espacios de características se usaron como entradas para un clasificador SVM binario clásico con matrices de kernel pre-calculadas. Para el proceso de entrenamiento se llevó a cabo una partición de 70-30 muestras, para entrenamiento se seleccionó el 70%, mientras que el 30% restante se usó para la validación, un proceso que fue repetido 10 veces. Los resultados obtenidos durante esta evaluación se presentan en la Tabla 2.

Tabla 2. Rendimiento promedio según un esquema de validación de 70–30% para la tarea de clasificación de RoIs con masas benignas y malignas de la misma mama (base de datos INBreast)

	INBreast		
	MLO-CC	MLO	CC
Sensibilidad	98.1	87.2	85.3
Especificidad	95.6	83.4	81.2
AUC	0.98	0.80	0.79
Accuracy	97.2	85.3	82.7

En este caso particular con la base de datos INBreast, los resultados reportados demuestran que la fusión de información de las dos proyecciones (MLO-CC) supera al desempeño logrado por cada una de las proyecciones por separado MLO y CC, respectivamente, alcanzando una sensibilidad del 98.1% frente 87.2% y 85.3% correspondiente a la proyección MLO y CC, mientras las especificidades alcanzada de 95.6 fue superior a 83.3% y 81.2% alcanzados por cada proyección por separado. Además, las métricas *Accuracy* y *AUC* muestran resultados notables, con 97.2 y 0.98 para esta base de datos. De igual manera que en el experimento anterior, los resultados de esta evaluación pueden ser visualizados mediante las curvas ROC, los cuales se ilustran en la figura 8.

4. Discusión

En este trabajo se ha propuesto un método novedoso que fusiona información visual completamente obtenida desde dos vistas diferentes de la misma masa para su clasificación entre benigna o maligna de acuerdo al criterio establecido en el sistema BI-RADS, comportándose como un modelo que se acerca al flujo de trabajo real de los radiólogos en escenarios clínicos reales, a la vez que cumple el requisito importante para ser implementado en cualquier sistema de soporte CAD. La importancia de esta estrategia se fundamenta en un par de descripciones complementarios multi-resolución, los mismos que se fusionan de manera óptima en un marco adaptable de aprendizaje basado en múltiples kernels.

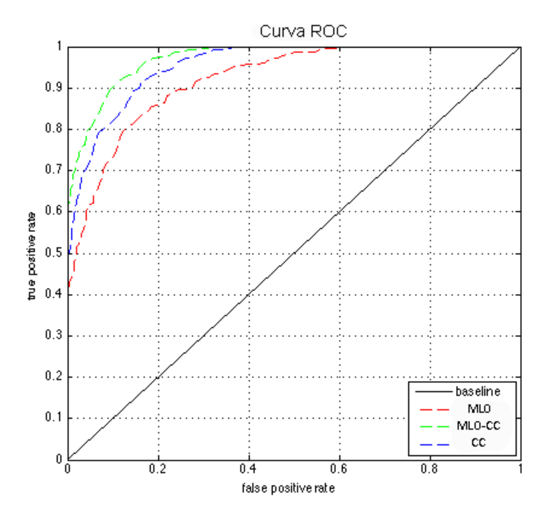


Figura 8. Curva ROC correspondiente al desempeño de la estrategia de fusión propuesta (MLO-CC) respecto a la clasificación usando información de una única vista MLO y CC, respectivamente

En este trabajo se aprovecharon las características visuales variables que pueden presentar entre una y otra proyección para refinar su clasificación, además de aprender del conocimiento implícito del radiólogo experto proporcionado sobre la anotación de la imágenes almacenadas en las bases de datos. La evaluación exhaustiva ha demostrado que el enfoque actual es altamente confiable para ser utilizado en escenarios clínicos reales y con cualquier estrategia de CAD para tareas de clasificación.

Frecuentemente, diferentes métodos para clasificar masas como benignas y malignas [15,16,14,17], segmentan la lesión y extraen diferentes características cuyo rendimiento para esta tarea de clasificación se ha evaluado mediante las curvas ROC o FROC. Olivier et. Alabama. [17] proporcionó una comparación cuantitativa de siete métodos que comparten algunas estrategias básicas, tales como, crecimiento de regiones para localizar la presencia de estructuras específicas, cluster o agrupamientos de píxeles y modelos de detección de contornos específicos que caracterizan lesiones potenciales. Este conjunto de métodos se evaluó con una parte de la base

de datos MIAS y una recopilación de datos privados, solo 261 casos que representan un conjunto de datos muy reducido, lo que reportaron áreas bajo la curva ROC de AUC=0.8, sin embargo, el número de falsos positivos por imagen alcanzó más de cinco por cada casos. Por el contrario, el esquema propuesto basado en fusión de vistas MLO-CC fue evaluado con una base de datos compuesta por RoIs, extraídos de la base de datos DDSM y INBreast obteniéndose valores del AUC de 0.98 y 0.95, respectivamente, en la tarea de clasificación como benigna y maligna. Así mismo por ejemplo, otros trabajos han utilizado diferentes análisis multi-resolución para construir un descriptor de bajo nivel. Tahmasbi et. al [18] han demostrado que una representación de forma útil, requiere una serie de momentos de Zernike [18, 19]. Así, evaluaron los momentos de Zernike calculados en un orden máximo (orden 17) y los dividieron en dos grupos, 32 de orden inferior y 32 de orden superior. La combinación de estos dos grupos resultó como el mejor descriptor de forma de las masas, reportando un AUC = 0.97 en la tarea de clasificación como benignas o malignas, para lo cual, usaron una red neuronal como clasificador mediante un algoritmo Perceptron de múltiples capas. Este método fue evaluado con la base de datos MIAS, específicamente usaron, 209 ROI normales, 67 ROI benignos y 54 ROI malignos y la conclusión fue que los momentos de Zernike de bajo orden mostraron poca capacidad de discriminación, reportando un AUC = 0.59, pero la combinación con órdenes superiores mejoró la resultados del método. En contraste, los resultados obtenidos en este trabajo superaron este método y fusionaron de manera óptima las representaciones de Zernike-Curvelet, con momentos de Zernike de orden de 45^{th} a 50^{th}. Este conjunto de momentos fue seleccionado debido a que observamos que la mayor parte de la energía de los coeficientes se concentraba en este rango y, además, el PSNR durante la reconstrucción de la imagen original era mayor de $40{:}dB$. De igual manera, Eltoukhy et. al [32] comparó las diferentes transformaciones wavelet y Curvelet para diferenciar cualquier tipo de lesión mamaria, específicamente, masas, calcificaciones, distorsión arquitectónica y asimetría bilateral. Para este propósito seleccionó un conjunto de 100 coeficientes de cada transformación y fueron usados con una estrategia de clasificación mediante los vecinos

cercanos K-NN implementado con una métrica Euclideana. Para la validación se usaron 322 mamogramas, 51 malignos, 64 benignos y 207 normales. Las imágenes fueron descompuestas en cuatro niveles diferentes con tres funciones de wavelets diferentes, entre estas: Daubechies-8, symlet y bi-ortogonal, mientras que se usaron cuatro niveles de descomposición para las curvelets. Los resultados demostraron una tasa de clasificación promedio para las 3 clases (benigna, maligna y normal) de 94.07% para las Curvelet y 90.05% para las wavelets. Posteriormente, estos autores mejoraron sus resultados [32], utilizando un esquema SVM y logrando un 96.56% con la misma base de datos. Sin embargo, el análisis fue realizado en una única proyección de mamografías y su base de datos es mucho más pequeña que la utilizada en esta investigación, lo que se puede observar que el rendimiento es similar. Asimismo, se han reportado diferentes técnicas estadísticas, estructurales y espectrales para analizar las mamografías [24], incluidas las matrices de coocurrencia [25]. Ramos et.al [25] presentaron una evaluación de los métodos basados en Ridgelet, Wavelet, Co-ocurrence y concluyeron que la wavelet superó a las otras representaciones, con un AUC de 0.9 cuando clasificaba las masas como benignas o malignas, resultados que han sido mejorados con el método propuesto en este trabajo.

Una gran ventaja del presente método es que, es completamente automático, las evaluaciones mostraron un rendimiento notable en diferentes bases de datos, clasificando masas en regiones bien definidas y algunas con una pobre calidad de la imagen. La estrategia se basa en la fusión óptima de la información visual de dos vistas de la misma lesión y proveída por dos descripciones complementarias cuyos pesos se pueden aprender de la base de datos. El tamaño RoI se puede considerar un parámetro importante en los resultados y su influencia no se evaluó en este trabajo, sin embargo, el hecho de que los descriptores, Zernike y Curvelet, sean por naturaleza una descomposición multi-resolución y multi-escala podrían mitigar los efectos causados por la selección de este parámetro. Otro limitación del método se debe a la falta de detección automática de la masa (RoIs) en las dos proyecciones, en este caso las regiones se proporcionan como una entrada. Sin embargo, esto podría considerarse un inconveniente, pero el objetivo de esta investigación fue

diseñar una estrategia integrable con un flujo de trabajo de radiólogo real y proporcionar alguna ayuda crucial al realizar tareas de diagnóstico. Se podría construir un método totalmente automático o de detección a partir de la presente estrategia mediante una simple búsqueda exhaustiva de toda la mamografía, pero esto estaba fuera del alcance de la presente investigación. Finalmente, el método muestra tiempos de ejecución asequibles del orden de 3-5 segundos.

5. Agradecimientos

Este trabajo fue parcialmente financiado por el Gobierno ecuatoriano a través de La Secretaría Nacional de Ciencia, Tecnología y Educación Superior (SENESCYT), Beca: 20110958 "convocatoria abierta 2011".

Referencias

[1] B. W. Stewart and C. P. Wild, "World cancer report 2014," 2014.

[2] M. Schneider, "Better detection: improving our chances, in: M.j. yaffe (ed.), digital mammography," in Digital Mammography, M. Yaffe, Ed., Fifth International Workshop on Digital Mammography. Toronto, Canada: Medical Physics Publishing, 2000, pp. 3–6.

[3] K. Kerlikowske, P. Carney, B. Geller, M. Mandelson, S. Taplin, K. Malvin,V. Ernster, N. Urban, G. Cutter, R. Rosenberg, and R. Ballard-Barbash, "Performance of screening mammography among women with and without a first-degree relative with breast cancer," Ann Intern Med.,vol. 133, no. 11, pp. 855–863, Dic 5 2000.

[4] M. Homer, Mammographic Interpretation: A Practical Approach, M.-H. S. Ed., Ed. New York, 1997.

[5] E. A. Sickles, C. J. DO´ rsi, L. W. Bassett, and E. al., ACR BI-RADS Mammography. In: ACR BI-RADS Atlas, Breast Imaging Reporting and Data System., fifth edit ed. American College of Radiology, Reston, VA, 2013.

[6] L. J. Grimm, S. V. Ghate, S. C. Yoon, C. M. Kuzmiak, C. Kim, and M. A. Mazurowski, "Predicting error in detecting mammographic masses among radiology trainees using statistical models based on BI-RADS features." Medical physics, vol. 41, no. 3, p. 31909, mar 2014. [Online]. Available: http://scitation.aip.org/content/aapm/journal/ medphys/41/3/10.1118/1.4866379

[7] R. M. Nishikawa, "Current status and future directions of computeraided diagnosis in mammography," Computerized Medical Imaging and Graphics, vol. 31, pp. 224–235, 2007.

[8] B. Zheng, C. Mello-Thoms, X. Wang, G. Abrams, J. Sumkin, D. Chough, M. Ganott, A. Lu, and D. Gur, "Interactive computer aided diagnosis of breast

masses: computerized selection of visually similar image sets from a reference library," Acad Radiol, vol. 14, pp. 917–927, 2007.
[9] D. Gur, J. Stalder, L. Hardesty, B. Zheng, J. Sumkin, D. Chough, B. Shindel, and H. Rockette, "Computer-aided detection performance in mammographic examination of masses: assessment." Radiology, vol. 233, pp. 418–423, 2004.
[10] L. A. L. Khoo, P. Taylor, and R. M. Given-Wilson, "Computer-aided detection in the united kingdom national breast screening programme: prospective study," Radiology., vol. 237(2), pp. 444–449, 2005.
[11] T. Freer and M. Ulissey, "Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center." Radiology, vol. 220, pp. 781–786, 2001.
[12] M. R. Rangayyana, F. J. Ayres, and J. L. Desautelsa, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," Journal of the Franklin Institute, vol. 344, pp. 312–348, 2007.
[13] N. Petrick, H. Chan, B. Sahiner, and D. Wei, "An adaptive density weighted contrast enhancement filter for mammographic breast mass detection," IEEE Trans. Med. Imaging, vol. 15, no. 1, pp. 59–67, 1996.
[14] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai, and H. Du, "Approaches for automated detection and classification of masses in mammograms," Pattern Recognition, vol. 39, pp. 646–668, 2006.
[15] A. Oliver, J. Freixenet, J. Mart´ı, E. P´erez, J. Pont, E. R. Denton, and R. Zwiggelaar, "A review of automatic mass detection and segmentation in mammographic images," Medical Image Analysis, vol. 14, pp. 87– 110, 2010.
[16] A. Tahmasbi, F. Saki, and S. B. Shokouhi, "Classification of benign and malignant masses based on zernike moments," Computers in Biology and Medicine, vol. 41, pp. 726–735, 2011.
[17] C.-H. Wei, S. Y. Chen, and X. Liu, "Mammogram retrieval on similar mass lesions," Computer Methods and programs in biomedicine, vol. 106, pp. 234–248, 2012.
[18] G. Papakostas, Y. Boutalis, D. Karras, and B. Mertzios, "A new class of zernike moments for computer vision applications," Information Sciences, vol. 177, no. 13, pp. 2802–2819, 2007.
[19] C. Wee and R. Paramesran, "On the computational aspects of Zernike moments," Image and Vision Computing, vol. 25, no. 6, pp. 967–980 , 2007.
[20] S. Li, M. C. Lee, and C. M. Pun, "Complex zernike moments features for shape-based image retrieval," IEEE Transactions on Systems, Man and Cybernetics, vol. 1, no. 39, pp. 227–237, 2009.
[21] C.-H. Wei, Y. Li, W. Chau, and C.-T. Li, "Trademark image retrieval: using synthetic features for describing global shape and interior structure of trademark," Pattern Recognition, vol. 42, no. 3, pp. 386–394, 2009.
[22] A. R. Dominguez and A. K. Nandi, "Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection," Computerized Medical Imaging and Graphics, vol. 32, pp. 304–315, 2008.

[23] R. P. Ramos, M. Z. do Nascimento, and D. C. Pereira, "Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms," Expert Systems with Applications, vol. 39, pp. 11 036–11 047, 2012.

[24] S. Nam and J. Choi, "A method of image enhancement and fractal dimension for detection of microcalcifications in mammogram," IEEE Engineering in Medicine and Biology Society, vol. 2, pp. 1009–1012, 1998.

[25] L. Bocchi, G. Coppini, J. Nori, and G. Valli, "Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks," Medical Engineering and Physics, vol. 26, pp. 303–312, 2004.

[26] S. Liu, C. Babbs, and E. Delp, "Multiresolution detection of speculated lesions in digital mammograms," IEEE Transactions on Image Processing, vol. 10, no. 6, pp. 874–884, 2001.

[27] M. Z. do Nascimento, A. S. Martins, L. A. Neves, R. P. Ramos, E. L. Flores, and G. A. Carrijo, "Classification of masses in mammographic image using wavelet domain features and polynomial classifier," Expert Systems with Applications, vol. 40, no. 15, pp. 6213–6221, 2013.

[28] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," Computerized Medical Imaging and Graphics, vol. 34, pp. 269–276, 2010.

[29] F. Moayedi, Z. Azimifar, R. Boostani, and S. Katebi, "Contourlet-based mammography mass classification using the svm family," Comput. Biol. Med., vol. 40, no. 4, pp. 373–383, 2010.

[30] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram," Computers in Biology and Medicine, vol. 40, pp. 384–391, 2010.

[31] F. Narvaez, G. Diaz, and E. Romero, "Multi-view information fusion for automatic bi-rads description of mammographic masses," in Medical Imaging 2011: Computer-Aided Diagnosis, vol. 7963. International Society for Optics and Photonics, 2011, p. 79630A.

[32] ——, "Automatic BI-RADS Description of Mammographic Masses," in Digital Mammography, vol. 6136, 2010, pp. 673–681.

[33] F. Narv´aez, G. D´ıaz, F. G´omez, and E. Romero, "A content-based retrieval of mammographic masses using the curvelet descriptor," in Proc. SPIE, vol. 8315, 2012, pp. 83 150A–83 150A–7.

[34] F. Narv´aez, G. D´ıaz, C. Poveda, and E. Romero, "An automatic BIRADS description of mammographic masses by fusing multiresolution features," Expert Systems with Applications, vol. 74, pp. 82–95, 2017.

[35] A. A. AbuBaker, R. S.Qahwaji, M. J. Aqel, and M. H. Saleh, "Mammogram image size reduction using 16-8 bit conversion technique," International Journal of Medical, Pharmaceutical Science and Engineering, vol. 1, no. 3, pp. 83–90, 2007.

[36] D. Kopans, Brast Imaging, Philadelphia, Ed. Lippicott-Raven, 1998.

[37] J. Beutel, H. Kundel, and R. V. Metter, Handbook of Medical Imaging, ser. Phys. Psychophys, S. Press, Ed. Bellingham, WA, 2000, vol. 1.
[38] A. C. of Radiology (ACR), "Illustrated breast imaging reporting and data system (bi-rads)," American College of Radiology, Reston, VA,, vol. third ed, p. 1998.
[39] C. Maggio, "State of the art of current modalities for the diagnosis of breast lesions," Eur. J Nucl Med Mol Imaging, vol. 31, no. 1, pp. S56–S69, 2004.
[40] S. X. Liao, "Image analysis with zernike moment descriptors," IEEE CCECE, pp. 700–703, 1997.
[41] K. M. Hosny, "Fast computation of accurate zernike moments," J Real-Time Image Proc, vol. 3, pp. 97–107, 2008.
[42] P. Dong and N. P. Galatsanos, "Affine transformation resistant watermarking based on image normalization," in Image Processing. 2002. Proceedings. 2002 International Conference on, vol. 3. IEEE, 2002, pp. 489–492.
[43] E. Candes, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," Multiscale Modeling and Simulation, vol. 5, no. 3, pp. 861–899, 2006.
[44] F. G´omez and E. Romero, "Texture characterization using a curvelet based descriptor," in Proceedings of the 14th Iberoamerican Conference on Pattern Recognition: Progress in Pattern Recognition, Image Analysis. Computer Vision, and Applications, 2009.
[45] M. Meselhy Eltoukhy, I. Faye, and B. Belhaouari Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation." Computers in biology and medicine, vol. 42, no. 1, pp. 123–128, jan 2012.
[46] G. Kom, A. Tiedeu, and M. Kom, "Automated detection of masses in mammograms by local adaptive thresholding," Computers in Biology and Medicine, vol. 37, pp. 37–48, 2007.
[47] S. Frintrop, G. Backer, and E. Rome, "Goal-directed search with a topdown modulated computational attention system," Pattern Recognit, pp.117–124, 2005.
[48] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 2005.
[49] M. Varma and B. Babu, "More generality in efficient multiple kernel learning," in ACM 26th Annu. Int. Conf. Mach. Learn, 2009, pp. 1065– 1072.
[50] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," IEEE Transactions on Neural Networks, vol. 13, no. 2, pp. 415–425, 2002.
[51] M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, "The digital database for screening mammography," in Proceedings of the Fifth International Workshop on Digital Mammography, Medical Physics Publishing, vol. M.J. Yaffe, ed, pp. 212–218, 2001.
[52] D. C. Moura, M. A. Guevara L´opez, and M. A. G. L´opez, "An evaluation of image descriptors combined with clinical data for breast cancer diagno-

sis." International journal of computer assisted radiology and surgery, vol. 8, no. 4, pp. 561–574, jul 2013

[53] A. Horsch, A. Hapfelmeier, and M. Elter, "Needs assessment for next generation computer-aided mammography reference image databases and evaluation studies," International Journal of Computer Assisted Radiology and Surgery, vol. 6, no. 6, pp. 749–767, 2011.

[54] X. Zhang, "A New Ensemble Learning Approach for Microcalcification Clusters Detection," Journal of Software, vol. 4, no. 9, pp. 1014–2021, 2009.

[55] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: toward a full-field digital mammographic database." Academic Radiology, vol. 19, no. 2, pp. 236–48, feb 2012.

[56] C.-C. Chang and C.-J. Lin., "LIBSVM : a library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 27, pp. 1–27, 2011.

Sobre el autor

Fabián Rodrigo Narváez Espinoza. (Y'77-M'05). Director de la Carrera de Biomedicina. Se formó como Ingeniero Electrónico en la Universidad Politécnica Salesiana (2006). Obtuvo la maestría en Ingeniería Biomédica en la Universidad Nacional de Colombia (2010), con distinción Meritoria, y posteriormente el Doctorado en Ingeniería Mecánica y Mecatrónica, en la misma universidad (2016), con distinción Meritoria. Se desempeñó durante 8 años como Ingeniero de desarrollo en el Centro de Telemedicina y el Departamento de Imágenes Diagnósticas de la Universidad Nacional de Colombia-Sede Bogotá, con la inclusión de tecnologías emergentes para la prestación de servicios de medicina a distancia bajo la modalidad de Telemedicina. Durante el 2011 al 2015 fue docente-investigador en la Universidad Central (Bogotá). Actualmente, se encuentra vinculado como docente - investigador en la Universidad Politécnica Salesiana, Sede-Quito, en la carrera de Ing. Mecatrónica. Sus áreas de interés se enfocan en el procesamiento de señales biomédicas, reconocimiento de patrones, aprendizaje de máquina, visión por computador, procesamiento de imágenes diagnósticas y sistemas de diagnóstico asistido por computador.

